GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2003, 06:13:09 ; Search time 51 Seconds
(without alignments)
487.075 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09724000/runat_16052003_145501_17219/app_query.fasta_1.263
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09724000_@CGN_1_1_40_@runat_16052003_145501_17219 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	85	19.2	1688	2	US-08-649-619B-2	Sequence 2, Appli
c	2	85	19.2	1979	2	US-08-649-619B-1	Sequence 1, Appli
c	3	85	19.2	2030	1	US-08-330-518-1	Sequence 1, Appli
c	4	85	19.2	2030	1	US-08-330-283-1	Sequence 1, Appli
c	5	85	19.2	2030	2	US-08-646-248-1	Sequence 1, Appli
c	6	85	19.2	2030	5	PCT-US95-13924-1	Sequence 1, Appli
c	7	85	19.2	2030	5	PCT-US95-13931-1	Sequence 1, Appli
c	8	78	17.6	1813	5	PCT-US94-12883-3	Sequence 3, Appli
c	9	78	17.6	1898	1	US-08-342-411A-1	Sequence 1, Appli
c	10	74.5	16.9	48974	4	US-08-920-422-17	Sequence 17, Appl
	11	73.5	16.6	1600	4	US-09-029-027B-1	Sequence 1, Appli
c	12	73.5	16.6	2542	4	US-09-305-681-5	Sequence 5, Appli
	13	72	16.3	1493	2	US-08-752-307B-6	Sequence 6, Appli
	14	72	16.3	1493	4	US-09-707-802-6	Sequence 6, Appli
	15	72	16.3	1493	4	US-09-991-326-6	Sequence 6, Appli
	16	72	16.3	2244	4	US-09-354-151-1	Sequence 1, Appli
	17	72	16.3	4086	1	US-08-313-181-1	Sequence 1, Appli
c	18	71.5	16.2	888	4	US-09-215-694-30	Sequence 30, Appl
	19	71.5	16.2	1736	4	US-09-162-524-2	Sequence 2, Appli
c	20	71.5	16.2	1798	4	US-09-305-681-1	Sequence 1, Appli
c	21	71.5	16.2	2416	1	US-08-592-126-61	Sequence 61, Appl
	22	71.5	16.2	31328	4	US-09-215-694-19	Sequence 19, Appl
	23	71	16.1	3807	2	US-08-816-755-1	Sequence 1, Appli
	24	71	16.1	3807	4	US-09-090-673-1	Sequence 1, Appli
	25	70.5	16.0	2219	1	US-08-606-322-1	Sequence 1, Appli
	26	70.5	16.0	3502	2	US-08-724-394A-16	Sequence 16, Appl
	27	70.5	16.0	4040	4	US-08-664-962B-1	Sequence 1, Appli
	28	70.5	16.0	4040	4	US-09-311-743-1	Sequence 1, Appli
	29	70	15.8	2743	1	US-08-396-479B-3	Sequence 3, Appli
	30	70	15.8	2743	1	US-08-818-823-3	Sequence 3, Appli
	31	70	15.8	2749	2	US-08-124-981A-1	Sequence 1, Appli
	32	70	15.8	2751	3	US-09-037-190-45	Sequence 45, Appl
	33	70	15.8	2751	3	US-09-037-192-45	Sequence 45, Appl
	34	70	15.8	2751	3	US-09-037-143-45	Sequence 45, Appl
	35	70	15.8	2751	4	US-09-049-691-45	Sequence 45, Appl
	36	70	15.8	2751	4	US-08-260-174-45	Sequence 45, Appl
	37	70	15.8	2751	4	US-09-338-128A-45	Sequence 45, Appl
	38	70	15.8	2751	4	US-09-037-192-45	Sequence 45, Appl
	39	70	15.8	2853	5	PCT-US94-07297-36	Sequence 36, Appl
	40	70	15.8	5502	4	US-09-232-346-45	Sequence 45, Appl
	41	69.5	15.7	1110	4	US-08-462-509B-3	Sequence 3, Appli
	42	69.5	15.7	1110	5	PCT-US95-05616-3	Sequence 3, Appli
	43	69.5	15.7	1654	3	US-08-991-426-3	Sequence 3, Appli
	44	69.5	15.7	1654	3	US-09-143-470-3	Sequence 3, Appli
c	45	69.5	15.7	2959	1	US-08-408-188A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-649-619B-2/c
; Sequence 2, Application US/08649619B
; Patent No. 5871916
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: : SAITO, HIROKO
; TITLE OF INVENTION: ECDN PROTEIN AND DNA
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLYNN, THIEL, BOUTELL &
; ADDRESSEE: TANIS, P.C.
; STREET: 2026 Rambling Road
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49008-1699
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC/XT/AT Compatible
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,619B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP6-226270
; FILING DATE: 21-SEPT-1994
; APPLICATION NUMBER: PCT/JP95/01909
; FILING DATE: 21-SEPT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Terryence F. Chapman
; REGISTRATION NUMBER: 32549

```
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1343
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1688
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homosapiens
US-08-649-619B-2

Alignment Scores:
Pred. No.:              1.65          Length:         1688
Score:                  85.00         Matches:        26
Percent Similarity:     47.76%        Conservative:   6
Best Local Similarity:  38.81%        Mismatches:     17
Query Match:            19.23%        Indels:         18
DB:                     2             Gaps:           4

US-09-724-000A-5 (1-81) x US-08-649-619B-2 (1-1688)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||      |||:::|||      |||:::|||||||||     :::
Db     692 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 633

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||      |||   |||   |||   :::   :::|||
Db     632 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 591

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                        |||||||||               |||||||||
Db     590 ---------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCGCCTTCCCCGGAGC 546

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     545 CCTGGCTGCCTGCCTCAGATC 525

RESULT 2
US-08-649-619B-1/c
; Sequence 1, Application US/08649619B
; Patent No. 5871916
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  : SAITO, HIROKO
;    TITLE OF INVENTION:  ECDN PROTEIN AND DNA
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL &
;      ADDRESSEE:  TANIS, P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/649,619B
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP6-226270
;      FILING DATE:  21-SEPT-1994
;      APPLICATION NUMBER:  PCT/JP95/01909
;      FILING DATE:  21-SEPT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1343
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1979
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homosapiens
;    IMMEDIATE SOURCE:
;      LIBRARY:  Human mammary gland cDNA
;      LIBRARY:  library
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  206..1591
;      IDENTIFICATION METHOD:  experimental examination
US-08-649-619B-1

Alignment Scores:
Pred. No.:              2.06          Length:         1979
Score:                  85.00         Matches:        26
Percent Similarity:     47.76%        Conservative:   6
Best Local Similarity:  38.81%        Mismatches:     17
Query Match:            19.23%        Indels:         18
DB:                     2             Gaps:           4

US-09-724-000A-5 (1-81) x US-08-649-619B-1 (1-1979)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||      |||:::|||      |||:::|||||||||     :::
Db     983 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 924

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||      |||   |||   |||   :::   :::|||
Db     923 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 882

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                        |||||||||               |||||||||
Db     881 ---------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCGCCTTCCCCGGAGC 837

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     836 CCTGGCTGCCTGCCTCAGATC 816

RESULT 3
US-08-330-518-1/c
; Sequence 1, Application US/08330518
; Patent No. 5607967
;  GENERAL INFORMATION:
;    APPLICANT:  Friedman, Eitan
;    APPLICANT:  Holloway, M. Katharine
;    APPLICANT:  Rodan, Gideon
;    APPLICANT:  Schmidt, Azriel
;    APPLICANT:  Vogel, Robert
;    TITLE OF INVENTION:  USE OF RECEPTOR POTENTIATORS
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  126 East Lincoln Avenue
;      CITY:  Rahway
;      STATE:  New Jersey
;      COUNTRY:  US
;      ZIP:  07065-0907
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/330,518
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dolan, Catherine A.
;      REGISTRATION NUMBER:  36,502
;      REFERENCE/DOCKET NUMBER:  19316
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 594-4283
;      TELEFAX:  (908) 594-4720
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2030 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-330-518-1

Alignment Scores:
Pred. No.:                2.13          Length:          2030
Score:                    85.00         Matches:         26
Percent Similarity:       47.76%        Conservative:    6
Best Local Similarity:    38.81%        Mismatches:      17
Query Match:              19.23%        Indels:          18
DB:                       1             Gaps:            4

US-09-724-000A-5 (1-81) x US-08-330-518-1 (1-2030)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||     |||:::|||      |||:::|||||||||      :::
Db    1022 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 963

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||     |||   |||   |||   :::   :::|||
Db     962 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 921

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                      |||||||||               |||||||||
Db     920 --------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCGCCTTCCCCGGAGC 876

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     875 CCTGGCTGCCTGCCTCAGATC 855

RESULT 4
US-08-330-283-1/c
; Sequence 1, Application US/08330283
; Patent No. 5679518
;  GENERAL INFORMATION:
;    APPLICANT:  Friedman, Eitan
;    APPLICANT:  Holloway, M. Katharine
;    APPLICANT:  Rodan, Gideon
;    APPLICANT:  Rutledge, Su Jane
;    APPLICANT:  Schmidt, Azriel
;    APPLICANT:  Vogel, Robert
;    TITLE OF INVENTION:  METHOD FOR FINDING RECEPTOR POTENTIATORS
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  126 East Lincoln Avenue
;      CITY:  Rahway
;      STATE:  New Jersey
;      COUNTRY:  US
;      ZIP:  07065-0907
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/330,283
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dolan, Catherine A.
;      REGISTRATION NUMBER:  36,502
;      REFERENCE/DOCKET NUMBER:  19327
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 594-4283
;      TELEFAX:  (908) 594-4720
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2030 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-330-283-1

Alignment Scores:
Pred. No.:                2.13          Length:          2030
Score:                    85.00         Matches:         26
Percent Similarity:       47.76%        Conservative:    6
Best Local Similarity:    38.81%        Mismatches:      17
Query Match:              19.23%        Indels:          18
DB:                       1             Gaps:            4

US-09-724-000A-5 (1-81) x US-08-330-283-1 (1-2030)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||     |||:::|||      |||:::|||||||||      :::
Db    1022 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 963

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||     |||   |||   |||   :::   :::|||
Db     962 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 921

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                      |||||||||               |||||||||
Db     920 --------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCGCCTTCCCCGGAGC 876

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     875 CCTGGCTGCCTGCCTCAGATC 855

RESULT 5
US-08-646-248-1/c
; Sequence 1, Application US/08646248
; Patent No. 5939322
;  GENERAL INFORMATION:
;    APPLICANT:  Friedman, Eitan
;    APPLICANT:  Holloway, M. Katharine
;    APPLICANT:  Rodan, Gideon
;    APPLICANT:  Rutledge, Su Jane
;    APPLICANT:  Schmidt, Azriel
;    APPLICANT:  Vogel, Robert
;    TITLE OF INVENTION:  METHOD FOR FINDING RECEPTOR POTENTIATORS
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  126 East Lincoln Avenue
;      CITY:  Rahway
;      STATE:  New Jersey
;      COUNTRY:  US
;      ZIP:  07065-0907
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/646,248
;       FILING DATE:  14-MAY-1996
;       CLASSIFICATION:  536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/330,283
;       FILING DATE:  27-OCT-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Dolan, Catherine A.
;       REGISTRATION NUMBER:  36,502
;       REFERENCE/DOCKET NUMBER:  19327
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (908) 594-4283
;       TELEFAX:  (908) 594-4720
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2030 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-646-248-1

Alignment Scores:
Pred. No.:                2.13          Length:          2030
Score:                    85.00         Matches:         26
Percent Similarity:       47.76%        Conservative:    6
Best Local Similarity:    38.81%        Mismatches:      17
Query Match:              19.23%        Indels:          18
DB:                       2             Gaps:            4

US-09-724-000A-5 (1-81) x US-08-646-248-1 (1-2030)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||.     |||:::|||      |||:::|||||||||      :::
Db    1022 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 963

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||      |||   |||   |||   :::   :::|||
Db     962 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 921

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                     |||||||||               |||||||||
Db     920 --------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCGCCTTCCCCGGAGC 876

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     875 CCTGGCTGCCTGCCTCAGATC 855

RESULT 6
PCT-US95-13924-1/c
; Sequence 1, Application PC/TUS9513924
;  GENERAL INFORMATION:
;     APPLICANT:  Friedman, Eitan
;     APPLICANT:  Holloway, M. Katharine
;     APPLICANT:  Rodan, Gideon
;     APPLICANT:  Rutledge, Su Jane
;     APPLICANT:  Schmidt, Azriel
;     APPLICANT:  Vogel, Robert
;     TITLE OF INVENTION:  METHOD FOR FINDING RECEPTOR POTENTIATORS
;     NUMBER OF SEQUENCES:  5
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Merck & Co., Inc.
;       STREET:  126 East Lincoln Avenue
;       CITY:  Rahway
;       STATE:  New Jersey
;       COUNTRY:  US
;       ZIP:  07065-0907
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/13924
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Quagliato, Carol S.
;       REGISTRATION NUMBER:  35,330
;       REFERENCE/DOCKET NUMBER:  19327 PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (908) 594-3809
;       TELEFAX:  (908) 594-4720
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2030 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
PCT-US95-13924-1

Alignment Scores:
Pred. No.:                2.13          Length:          2030
Score:                    85.00         Matches:         26
Percent Similarity:       47.76%        Conservative:    6
Best Local Similarity:    38.81%        Mismatches:      17
Query Match:              19.23%        Indels:          18
DB:                       5             Gaps:            4

US-09-724-000A-5 (1-81) x PCT-US95-13924-1 (1-2030)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||      |||:::|||      |||:::|||||||||      :::
Db    1022 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 963

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||      |||   |||   |||   :::   :::|||
Db     962 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 921

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                     |||||||||               |||||||||
Db     920 --------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCGCCTTCCCCGGAGC 876

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     875 CCTGGCTGCCTGCCTCAGATC 855

RESULT 7
PCT-US95-13931-1/c
; Sequence 1, Application PC/TUS9513931
;  GENERAL INFORMATION:
;     APPLICANT:  Friedman, Eitan
;     APPLICANT:  Holloway, M. Katharine
;     APPLICANT:  Rodan, Gideon
;     APPLICANT:  Schmidt, Azriel
;     APPLICANT:  Vogel, Robert
;     TITLE OF INVENTION:  USE OF RECEPTOR POTENTIATORS
;     NUMBER OF SEQUENCES:  5
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Merck & Co., Inc.
;       STREET:  126 East Lincoln Avenue
;       CITY:  Rahway
;       STATE:  New Jersey
;       COUNTRY:  US
;       ZIP:  07065-0907
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/13931
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Quagliato, Carol S.
;       REGISTRATION NUMBER:  35,330
;       REFERENCE/DOCKET NUMBER:  19316 PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (908) 594-3809
;       TELEFAX:  (908) 594-4720
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2030 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
PCT-US95-13931-1

Alignment Scores:
Pred. No.:              2.13          Length:         2030
Score:                  85.00         Matches:        26
Percent Similarity:     47.76%        Conservative:   6
Best Local Similarity:  38.81%        Mismatches:     17
Query Match:            19.23%        Indels:         18
DB:                     5             Gaps:           4

US-09-724-000A-5 (1-81) x PCT-US95-13931-1 (1-2030)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||       |||:::|||      |||:::|||||||||       :::
Db    1022 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 963

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||      |||   |||   |||   :::   :::|||
Db     962 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 921

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                      |||||||||               |||||||||
Db     920 --------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCGCCTTCCCCGGAGC 876

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     875 CCTGGCTGCCTGCCTCAGATC 855

RESULT 8
PCT-US94-12883-3/c
; Sequence 3, Application PC/TUS9412883
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  UBIQUITOUS NUCLEAR RECEPTOR: COMPOSITIONS AND
;     TITLE OF INVENTION:  METHODS
;     NUMBER OF SEQUENCES:  38
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Arnold, White & Durkee
;       STREET:  P.O. Box 4433
;       CITY:  Houston
;       STATE:  Texas
;       COUNTRY:  United States of America
;       ZIP:  77210
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS/ASCII
;       SOFTWARE:  PatentIn Release #1.0, Version
;       SOFTWARE:  #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/12883
;       FILING DATE:  Concurrently Herewith
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/152,003
;       FILING DATE:  10-NOV-1993
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BARBARA S. KITCHELL
;       REGISTRATION NUMBER:  33,928
;       REFERENCE/DOCKET NUMBER:  ARCD154P--
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (512) 418-3000
;       TELEFAX:  (713) 789-2679
;       TELEX:  79-0924
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1813 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
PCT-US94-12883-3

Alignment Scores:
Pred. No.:              11.7          Length:         1813
Score:                  78.00         Matches:        25
Percent Similarity:     46.27%        Conservative:   6
Best Local Similarity:  37.31%        Mismatches:     18
Query Match:            17.65%        Indels:         18
DB:                     5             Gaps:           4

US-09-724-000A-5 (1-81) x PCT-US94-12883-3 (1-1813)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||       |||:::|||      |||:::|||||||||       :::
Db     760 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 701

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||      |||   |||   |||   :::   :::|||
Db     700 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 659

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                      |||||||||               |||   |||
Db     658 --------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCTCCTTCCCCGGAGC 614

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     613 CCTGGCTGCCTGCCTCAGATC 593

RESULT 9
US-08-342-411A-1/c
; Sequence 1, Application US/08342411A
; Patent No. 5639616
;  GENERAL INFORMATION:
;     APPLICANT:  LIAO, Shutsung
;     APPLICANT:  SONG, Ching
;     TITLE OF INVENTION:  UBIQUITOUS NUCLEAR RECEPTOR:
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS
;     NUMBER OF SEQUENCES:  38
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Arnold, White & Durkee
;       STREET:  P.O. Box 4433
;       CITY:  Houston
;       STATE:  TX
;       COUNTRY:  USA
;       ZIP:  77210-4433
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/342,411A
;      FILING DATE:  18-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KITCHELL, BARBARA S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  ARCD154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;      TELEX:  79-0924
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1898 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  71..1450
US-08-342-411A-1

Alignment Scores:
Pred. No.:              12.4          Length:        1898
Score:                  78.00         Matches:       25
Percent Similarity:     46.27%        Conservative:  6
Best Local Similarity:  37.31%        Mismatches:    18
Query Match:            17.65%        Indels:        18
DB:                     1             Gaps:          4

US-09-724-000A-5 (1-81) x US-08-342-411A-1 (1-1898)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||      |||:::|||      |||:::|||||||||      :::
Db     845 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 786

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||      |||   |||   |||   :::   :::|||
Db     785 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 744

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                     |||||||||             |||   |||
Db     743 --------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCTCCTTCCCCGGAGC 699

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     698 CCTGGCTGCCTGCCTCAGATC 678

RESULT 10
US-08-920-422-17/c
; Sequence 17, Application US/08920422A
; Patent No. 6255473
; GENERAL INFORMATION:
;  APPLICANT: Vitek, Michael P.
;  APPLICANT: Mitsuda, No. 6255473iaki
;  APPLICANT: Roses, Allen D.
;  TITLE OF INVENTION: Presenilin-1 Gene Promoter
;  FILE REFERENCE: VITEKPRESENILIN
;  CURRENT APPLICATION NUMBER: US/08/920,422A
;  CURRENT FILING DATE: 1997-08-29
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 48974
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-08-920-422-17

Alignment Scores:
Pred. No.:              2.88e+03      Length:        48974
Score:                  74.50         Matches:       24
Percent Similarity:     45.00%        Conservative:  12
Best Local Similarity:  30.00%        Mismatches:    32
Query Match:            16.86%        Indels:        13
DB:                     4             Gaps:          3

US-09-724-000A-5 (1-81) x US-08-920-422-17 (1-48974)

Qy       4 LeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePheSerThrGlu 23
           |||   :::   :::   :::|||:::   :::      |||:::      |||
Db   23314 CTCTCCATTCCCAATGCTGTGTGTGTGTGTGTGTGTGTGTTTGCATGTCTGTCTGTCTG- 23256

Qy      24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArg 43
                  ||||||   |||                    :::|||||||||
Db   23255 TCTCTATTACTTCCTGCTTCAGCC-----------ACTCTGCCTGTCTGCTGCCATAAC 23208

Qy      44 ValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCysLysProCys 63
             |||                 ::::::   |||         ||||||||   ||||||
Db   23207 TTCCCT--------------GACATGATGGGTGATGGTGATCGTTTATGCTGCCCCTGT 23163

Qy      64 LysLeuGluProGlu---------ProArgLeuTrpValValProGlyAlaLeuProGln 80
           |||   ||||||:::         |||   :::   :::      |||   |||   |||
Db   23162 AAGAATGAACCCAAATAAAGCCCTCCTATAATTTGCCTTGGTCATGGTGTTTTATCACAG 23103

RESULT 11
US-09-029-027B-1
; Sequence 1, Application US/09029027B
; Patent No. 6441133
; GENERAL INFORMATION:
;  APPLICANT: Walker, Philippe
;  TITLE OF INVENTION: No. 6441133el TRH Receptor
;  FILE REFERENCE: Walker appl.
;  CURRENT APPLICATION NUMBER: US/09/029,027B
;  CURRENT FILING DATE:  1998-01-27
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1600
;   TYPE: DNA
;   ORGANISM: Rat
US-09-029-027B-1

Alignment Scores:
Pred. No.:              32.4          Length:        1600
Score:                  73.50         Matches:       17
Percent Similarity:     50.98%        Conservative:  9
Best Local Similarity:  33.33%        Mismatches:    20
Query Match:            16.63%        Indels:        5
DB:                     4             Gaps:          2

US-09-724-000A-5 (1-81) x US-09-029-027B-1 (1-1600)

Qy      33 SerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsnSerThrAsnLeu 52
           :::|||             ||||||          |||   |||           :::
Db   1168 ACTGGTACTGCTCAACTCCTTTGTGGCCCAGCCTTTCCTGGACCCCTGGGTCCTGCTGTT 1227

Qy      53 LysGlyHis---------HisValArgLeuCysLysProCysLysLeuGluProGluPro 69
             |||         |||   :::   |||:::||||||:::|||:::|||:::
Db   1228 CTGCCGCACCTGTGTCTACACCAACAGCGCTGTCAACCCTGTCGTCTACAGCCTGAT--- 1284

Qy      70 ArgLeuTrpValValProGlyAlaLeuProGln 80
              :::      |||||||||   ||||||:::
Db   1285 ---GTCACAGAAGTTCCGGGCGGCCTTCCTGAA 1314

RESULT 12
US-09-305-681-5/c
; Sequence 5, Application US/09305681
; Patent No. 6406899
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Highly active alkaline phosphatase
;    NUMBER OF SEQUENCES:  54
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/305,681
;      FILING DATE:
;      CLASSIFICATION:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2542 base pairs
;      TYPE:  nucleotide
;      STRANDEDNESS:  single strand
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
US-09-305-681-5

Alignment Scores:
Pred. No.:              61.6          Length:        2542
Score:                  73.50         Matches:       22
Percent Similarity:     44.64%        Conservative:  3
Best Local Similarity:  39.29%        Mismatches:    22
Query Match:            16.63%        Indels:        9
DB:                     4             Gaps:          4

US-09-724-000A-5 (1-81) x US-09-305-681-5 (1-2542)

Qy      24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCys------ 41
           |||  |||  ||||||          |||        ||||||||||||
Db    1400 GGGGCACGGCCGCCTGCTGCCGGTACGAGGGGT------CCTCGCTTATGCTGTCATTAA 1347

Qy      42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArg------Leu 59
           ||||||     ||||||:::           :::||||||     |||
Db    1346 CATCGGGCCTTAAGCCCCCACCAAGCACGTAGCCAGGGCCATTGCCATAGAGGATGGAGG 1287

Qy      60 CysLysProCysLysLeuGluPro---------GluProArgLeuTrp 72
           |||:::  |||  |||  |||             |||  |||
Db    1286 TGTAGGACTTGTTGTCTGAGGCCTTGCTGGGGGCCAGACCGAAAATGG 1239

RESULT 13
US-08-752-307B-6
; Sequence 6, Application US/08752307B
; Patent No. 5952171
;  GENERAL INFORMATION:
;    APPLICANT:  McCarthy, Sean A.
;    APPLICANT:  Gearing, David P.
;    APPLICANT:  Levinson, Douglas A.
;    TITLE OF INVENTION:  METHOD FOR IDENTIFYING GENES
;    TITLE OF INVENTION:  ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/752,307B
;      FILING DATE:  19-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  09404/020001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1493 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  99...1493
US-08-752-307B-6

Alignment Scores:
Pred. No.:              43.8          Length:        1493
Score:                  72.00         Matches:       18
Percent Similarity:     39.39%        Conservative:  8
Best Local Similarity:  27.27%        Mismatches:    24
Query Match:            16.29%        Indels:        16
DB:                     2             Gaps:          3

US-09-724-000A-5 (1-81) x US-08-752-307B-6 (1-1493)

Qy      27 ArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCys--------------- 41
           ||||||:::  :::     |||     ||||||  |||
Db    1174 AGGCCATCTCCATCC---GTGGGCTCAGCAACGAGACGCTGCTCATCACCTCGGCCCAGA 1230

Qy      42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHis----------------- 55
           :::|||  |||  |||  :::|||
Db    1231 AGAGCCATTCCGGGGCCTACCAGTGCTTCGCTACCCGCAAGGCCCAGACCGCCCAGGACT 1290

Qy      56 ----------HisValArgLeuCysLysProCysLysLeuGluProGluProArgLeu 71
                    |||:::|||:::  :::|||           |||     |||
Db    1291 TTGCCATCATTGCACTTGAGGATGGCACGCCCCGCATCGTCTCGTCCTTCAGCGAGAAGG 1350

Qy      72 TrpValValProGlyAla 77
           |||     ||||||:::
Db    1351 TGGTCAACCCCGGGGAGC 1368

RESULT 14
US-09-707-802-6
; Sequence 6, Application US/09707802
; Patent No. 6391586
;  GENERAL INFORMATION:
;       APPLICANT: McCarthy, Sean A.
;                  Gearing, David P.
;                  Levinson, Douglas A.
;       TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
;                           ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
;       NUMBER OF SEQUENCES: 14
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fish & Richardson, P.C.
;            STREET: 225 Franklin Street
;            CITY: Boston
;            STATE: MA
;            COUNTRY: US
;            ZIP: 02110-2804
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: Windows95
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/707,802
;            FILING DATE: 07-No. 6391586-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/752,307
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Meiklejohn, Ph.D., Anita L.
;            REGISTRATION NUMBER: 35,283
;            REFERENCE/DOCKET NUMBER: 09404/020001
```

```
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 617-542-5070
;              TELEFAX: 617-542-8906
;              TELEX: 200154
;    INFORMATION FOR SEQ ID NO: 6:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 1493 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         FEATURE:
;              NAME/KEY:  Coding Sequence
;              LOCATION:  99...1493
;         SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-707-802-6

Alignment Scores:
Pred. No.:              43.8          Length:          1493
Score:                  72.00         Matches:         18
Percent Similarity:     39.39%        Conservative:    8
Best Local Similarity:  27.27%        Mismatches:      24
Query Match:            16.29%        Indels:          16
DB:                     4             Gaps:            3

US-09-724-000A-5 (1-81) x US-09-707-802-6 (1-1493)

Qy      27 ArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCys--------------- 41
           ||||||:::   :::      |||      ||||||   |||
Db    1174 AGGCCATCTCCATCC---GTGGGCTCAGCAACGAGACGCTGCTCATCACCTCGGCCCAGA 1230

Qy      42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHis------------------ 55
             :::|||   |||   |||   :::|||
Db    1231 AGAGCCATTCCGGGGCCTACCAGTGCTTCGCTACCCGCAAGGCCCAGACCGCCCAGGACT 1290

Qy      56 ------------HisValArgLeuCysLysProCysLysLeuGluProGluProArgLeu 71
                      |||:::|||:::   :::|||          |||      |||
Db    1291 TTGCCATCATTGCACTTGAGGATGGCACGCCCCGCATCGTCTCGTCCTTCAGCGAGAAGG 1350

Qy      72 TrpValValProGlyAla 77
           |||      ||||||:::
Db    1351 TGGTCAACCCCGGGGAGC 1368

RESULT 15
US-09-991-326-6
; Sequence 6, Application US/09991326
; Patent No. 6395872
;   GENERAL INFORMATION:
;         APPLICANT: McCarthy, Sean A.
;                    Gearing, David P.
;                    Levinson, Douglas A.
;         TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
;                             ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
;         NUMBER OF SEQUENCES: 14
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Fish & Richardson, P.C.
;              STREET: 225 Franklin Street
;              CITY: Boston
;              STATE: MA
;              COUNTRY: US
;              ZIP: 02110-2804
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Diskette
;              COMPUTER: IBM Compatible
;              OPERATING SYSTEM: Windows95
;              SOFTWARE: FastSEQ for Windows Version 2.0
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/991,326
;              FILING DATE: 21-No. 6395872-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/752,307
;              FILING DATE: 19-NOV-1996
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Meiklejohn, Ph.D., Anita L.
;              REGISTRATION NUMBER: 35,283
;              REFERENCE/DOCKET NUMBER: 09404/020002
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 617-542-5070
;              TELEFAX: 617-542-8906
;              TELEX: 200154
;    INFORMATION FOR SEQ ID NO: 6:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 1493 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         FEATURE:
;              NAME/KEY:  Coding Sequence
;              LOCATION:  99...1493
;         SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-991-326-6

Alignment Scores:
Pred. No.:              43.8          Length:          1493
Score:                  72.00         Matches:         18
Percent Similarity:     39.39%        Conservative:    8
Best Local Similarity:  27.27%        Mismatches:      24
Query Match:            16.29%        Indels:          16
DB:                     4             Gaps:            3

US-09-724-000A-5 (1-81) x US-09-991-326-6 (1-1493)

Qy      27 ArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCys--------------- 41
           ||||||:::   :::      |||      ||||||   |||
Db    1174 AGGCCATCTCCATCC---GTGGGCTCAGCAACGAGACGCTGCTCATCACCTCGGCCCAGA 1230

Qy      42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHis------------------ 55
             :::|||   |||   |||   :::|||
Db    1231 AGAGCCATTCCGGGGCCTACCAGTGCTTCGCTACCCGCAAGGCCCAGACCGCCCAGGACT 1290

Qy      56 ------------HisValArgLeuCysLysProCysLysLeuGluProGluProArgLeu 71
                      |||:::|||:::   :::|||          |||      |||
Db    1291 TTGCCATCATTGCACTTGAGGATGGCACGCCCCGCATCGTCTCGTCCTTCAGCGAGAAGG 1350

Qy      72 TrpValValProGlyAla 77
           |||      ||||||:::
Db    1351 TGGTCAACCCCGGGGAGC 1368


Search completed: May 17, 2003, 13:36:54
Job time : 65 secs
```

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2003, 11:01:51 ; Search time 147 Seconds
(without alignments)
710.782 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 810007 seqs, 644969091 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09724000/runat_16052003_145502_17248/app_query.fasta_1.263
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09724000_@CGN_1_1_135_@runat_16052003_145502_17248
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	442	100.0	342	9	US-10-066-543-3006	Sequence 3006, Ap
c 2	442	100.0	379	9	US-10-046-935-671	Sequence 671, App
c 3	442	100.0	379	9	US-09-878-178-671	Sequence 671, App
c 4	442	100.0	379	9	US-10-146-502-671	Sequence 671, App
5	442	100.0	800	9	US-09-981-353-177	Sequence 177, App
6	442	100.0	804	9	US-10-028-072-149	Sequence 149, App
7	442	100.0	804	9	US-10-121-049-149	Sequence 149, App
8	442	100.0	804	9	US-10-123-904-149	Sequence 149, App
9	442	100.0	804	9	US-10-140-470-149	Sequence 149, App
10	442	100.0	804	9	US-10-175-746-149	Sequence 149, App
11	442	100.0	804	9	US-10-176-918-149	Sequence 149, App
12	442	100.0	804	9	US-10-176-921-149	Sequence 149, App
13	442	100.0	804	9	US-10-137-865-149	Sequence 149, App
14	442	100.0	804	9	US-10-140-474-149	Sequence 149, App
15	442	100.0	804	9	US-10-142-431-149	Sequence 149, App
16	442	100.0	804	9	US-10-143-114-149	Sequence 149, App
17	442	100.0	804	9	US-10-140-002-149	Sequence 149, App
18	442	100.0	804	9	US-10-142-419-149	Sequence 149, App
19	442	100.0	804	9	US-10-123-262-149	Sequence 149, App
20	442	100.0	804	9	US-10-142-423-149	Sequence 149, App
21	442	100.0	804	9	US-10-121-050-149	Sequence 149, App
22	442	100.0	804	9	US-10-141-755-149	Sequence 149, App
23	442	100.0	804	9	US-10-143-032-149	Sequence 149, App
24	442	100.0	804	9	US-10-123-108-149	Sequence 149, App
25	442	100.0	804	9	US-10-123-236-149	Sequence 149, App
26	442	100.0	804	9	US-10-123-261-149	Sequence 149, App
27	442	100.0	804	9	US-10-140-921-149	Sequence 149, App
28	442	100.0	804	9	US-10-140-928-149	Sequence 149, App
29	442	100.0	804	9	US-10-121-045-149	Sequence 149, App
30	442	100.0	804	9	US-10-123-292-149	Sequence 149, App
31	442	100.0	804	9	US-10-123-903-149	Sequence 149, App
32	442	100.0	804	9	US-10-124-819-149	Sequence 149, App
33	442	100.0	804	9	US-10-124-822-149	Sequence 149, App
34	442	100.0	804	9	US-10-140-925-149	Sequence 149, App
35	442	100.0	804	9	US-10-160-498-149	Sequence 149, App
36	442	100.0	804	9	US-10-121-041-149	Sequence 149, App
37	442	100.0	804	9	US-10-121-043-149	Sequence 149, App
38	442	100.0	804	9	US-10-121-047-149	Sequence 149, App
39	442	100.0	804	9	US-10-123-215-149	Sequence 149, App
40	442	100.0	804	9	US-10-123-902-149	Sequence 149, App
41	442	100.0	804	9	US-10-123-908-149	Sequence 149, App
42	442	100.0	804	9	US-10-123-909-149	Sequence 149, App
43	442	100.0	804	9	US-10-123-910-149	Sequence 149, App
44	442	100.0	804	9	US-10-124-813-149	Sequence 149, App
45	442	100.0	804	9	US-10-124-817-149	Sequence 149, App

ALIGNMENTS

RESULT 1
US-10-066-543-3006/c
; Sequence 3006, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3006
; LENGTH: 342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-3006

```
Alignment Scores:
Pred. No.:                2.48e-46        Length:         342
Score:                    442.00          Matches:        81
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       9               Gaps:           0

US-09-724-000A-5 (1-81) x US-10-066-543-3006 (1-342)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 242

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 182

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 122

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 62

Qy      81 Val 81
           |||
Db      61 GTG 59

RESULT 2
US-10-046-935-671/c
; Sequence 671, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 671
;   LENGTH: 379
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 20, 21
;   OTHER INFORMATION: n = A,T,C or G
US-10-046-935-671

Alignment Scores:
Pred. No.:                2.81e-46        Length:         379
Score:                    442.00          Matches:        81
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       9               Gaps:           0

US-09-724-000A-5 (1-81) x US-10-046-935-671 (1-379)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 242

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 182

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 122

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 62

Qy      81 Val 81
           |||
Db      61 GTG 59

RESULT 3
US-09-878-178-671/c
; Sequence 671, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527
;  CURRENT APPLICATION NUMBER: US/09/878,178
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 2237
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 671
;   LENGTH: 379
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(379)
;   OTHER INFORMATION: n = A,T,C or G
US-09-878-178-671

Alignment Scores:
Pred. No.:                2.81e-46        Length:         379
Score:                    442.00          Matches:        81
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       9               Gaps:           0

US-09-724-000A-5 (1-81) x US-09-878-178-671 (1-379)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 242

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 182

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 122

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 62

Qy      81 Val 81
           |||
Db      61 GTG 59

RESULT 4
US-10-146-502-671/c
; Sequence 671, Application US/10146502
```

```
; Publication No. US20030069180A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C2
;  CURRENT APPLICATION NUMBER: US/10/146,502
;  CURRENT FILING DATE:  2002-05-14
;  NUMBER OF SEQ ID NOS: 2241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 671
;   LENGTH: 379
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 20, 21
;   OTHER INFORMATION: n = A,T,C or G
US-10-146-502-671

Alignment Scores:
Pred. No.:              2.81e-46        Length:         379
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-724-000A-5 (1-81) x US-10-146-502-671 (1-379)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 242

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 182

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 122

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 62

Qy     81 Val 81
          |||
Db     61 GTG 59

RESULT 5
US-09-981-353-177
; Sequence 177, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 177
;   LENGTH: 800
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020160382A1 1736965CB1
US-09-981-353-177

Alignment Scores:
Pred. No.:              7.13e-46        Length:         800
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-724-000A-5 (1-81) x US-09-981-353-177 (1-800)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     25 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 84

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 144

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 204

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 264

Qy     81 Val 81
          |||
Db    265 GTG 267

RESULT 6
US-10-028-072-149
; Sequence 149, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang
;  TITLE OF INVENTION:
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/028,072
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/049911
;  PRIOR FILING DATE: 1997-06-18
;  PRIOR APPLICATION NUMBER: 60/056974
;  PRIOR FILING DATE: 1997-08-26
;  PRIOR APPLICATION NUMBER: 60/059113
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059115
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059117
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059122
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059184
;  PRIOR FILING DATE: 1997-09-17
```

; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 19/98-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19

```
;  PRIOR APPLICATION NUMBER: 60/090349
;  PRIOR FILING DATE: 1998-06-23
;  PRIOR APPLICATION NUMBER: 60/090429
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090445
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090538
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090863
;  PRIOR FILING DATE: 1998-06-26
;  PRIOR APPLICATION NUMBER: 60/091360
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091519
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091982
;  PRIOR FILING DATE: 1998-07-07

Alignment Scores:
Pred. No.:              7.18e-46        Length:         804
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-724-000A-5 (1-81) x US-10-028-072-149 (1-804)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy      81 Val 81
           |||
Db     282 GTG 284

RESULT 7
US-10-121-049-149
; Sequence 149, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C17
;  CURRENT APPLICATION NUMBER: US/10/121,049
;  CURRENT FILING DATE:  2002-04-12
;       Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-121-049-149

Alignment Scores:
Pred. No.:              7.18e-46        Length:         804
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-724-000A-5 (1-81) x US-10-121-049-149 (1-804)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy      81 Val 81
           |||
Db     282 GTG 284

RESULT 8
US-10-123-904-149
; Sequence 149, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C54
;  CURRENT APPLICATION NUMBER: US/10/123,904
;  CURRENT FILING DATE:  2002-04-16
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-123-904-149
```

```
Alignment Scores:
Pred. No.:                7.18e-46        Length:         804
Score:                    442.00          Matches:        81
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       9               Gaps:           0

US-09-724-000A-5 (1-81) x US-10-123-904-149 (1-804)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy      81 Val 81
           |||
Db     282 GTG 284

RESULT 9
US-10-140-470-149
; Sequence 149, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C160
;  CURRENT APPLICATION NUMBER: US/10/140,470
;  CURRENT FILING DATE:  2002-05-06
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-140-470-149

Alignment Scores:
Pred. No.:                7.18e-46        Length:         804
Score:                    442.00          Matches:        81
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       9               Gaps:           0

US-09-724-000A-5 (1-81) x US-10-140-470-149 (1-804)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy      81 Val 81
           |||
Db     282 GTG 284

RESULT 10
US-10-175-746-149
; Sequence 149, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C353
;  CURRENT APPLICATION NUMBER: US/10/175,746
;  CURRENT FILING DATE:  2002-06-19
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-175-746-149

Alignment Scores:
Pred. No.:                7.18e-46        Length:         804
Score:                    442.00          Matches:        81
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       9               Gaps:           0

US-09-724-000A-5 (1-81) x US-10-175-746-149 (1-804)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
```

```
Db    102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
Db    162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
Db    222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy     81 Val 81
Db    282 GTG 284
```

RESULT 11
US-10-176-918-149
; Sequence 149, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-918-149

Alignment Scores:

Pred. No.:	7.18e-46	Length:	804
Score:	442.00	Matches:	81
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-724-000A-5 (1-81) x US-10-176-918-149 (1-804)

```
Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db     42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db    102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
Db    162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
Db    222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy     81 Val 81
Db    282 GTG 284
```

RESULT 12
US-10-176-921-149
; Sequence 149, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C288
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-921-149

Alignment Scores:

Pred. No.:	7.18e-46	Length:	804
Score:	442.00	Matches:	81
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-724-000A-5 (1-81) x US-10-176-921-149 (1-804)

```
Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db     42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db    102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
Db    162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
Db    222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy     81 Val 81
Db    282 GTG 284
```

RESULT 13

```
US-10-137-865-149
; Sequence 149, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C154
;  CURRENT APPLICATION NUMBER: US/10/137,865
;  CURRENT FILING DATE:  2002-05-03
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-137-865-149

Alignment Scores:
Pred. No.:              7.18e-46      Length:          804
Score:                  442.00        Matches:         81
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     9             Gaps:            0

US-09-724-000A-5 (1-81) x US-10-137-865-149 (1-804)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy     81 Val 81
          |||
Db    282 GTG 284

RESULT 14
US-10-140-474-149
; Sequence 149, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C162
;  CURRENT APPLICATION NUMBER: US/10/140,474
;  CURRENT FILING DATE:  2002-05-06
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-140-474-149

Alignment Scores:
Pred. No.:              7.18e-46      Length:          804
Score:                  442.00        Matches:         81
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     9             Gaps:            0

US-09-724-000A-5 (1-81) x US-10-140-474-149 (1-804)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy     81 Val 81
          |||
Db    282 GTG 284

RESULT 15
US-10-142-431-149
; Sequence 149, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
```

```
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C251
;  CURRENT APPLICATION NUMBER: US/10/142,431
;  CURRENT FILING DATE:  2002-05-10
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-142-431-149

Alignment Scores:
Pred. No.:              7.18e-46        Length:         804
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-724-000A-5 (1-81) x US-10-142-431-149 (1-804)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy     81 Val 81
          |||
Db    282 GTG 284


Search completed: May 17, 2003, 14:37:03
Job time : 152 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2003, 03:25:14 ; Search time 159 Seconds
(without alignments)
1147.244 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09724000/runat_16052003_145459_17136/app_query.fasta_1.263
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09724000_@CGN_1_1_263_@runat_16052003_145459_17136 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	442	100.0	382	24	ABL37082	Human colon tumour
2	442	100.0	485	24	ABK27569	Human colon cancer
3	442	100.0	804	22	AAS21318	Human cDNA sequenc
4	442	100.0	806	24	AAD27025	Human Secreted epi
5	442	100.0	908	22	AAF97898	Human secreted pro
c 6	400	90.5	797	21	AAA72230	Human CASB gene pa
c 7	400	90.5	801	21	AAA72224	Human CASB611 cDNA
8	228.5	51.7	744	24	AAD27024	Mouse Secreted epi
9	178	40.3	4159	24	AAD27026	Human Secreted epi
10	92	20.8	487	22	AAS39569	cDNA encoding nove
11	92	20.8	487	22	AAK88908	Human digestive sy
c 12	91	20.6	619	23	ABK41859	cDNA encoding nove
13	86.5	19.6	2167	15	AAQ57015	PKC gamma. Bos ta
c 14	85	19.2	1622	21	AAC98948	Human pancreatic c
c 15	85	19.2	1688	17	AAT27617	Steroid hormone re
c 16	85	19.2	1979	17	AAT27616	Human foetal lung
c 17	85	19.2	2010	24	ABK84201	Human cDNA differe
c 18	85	19.2	2030	15	AAQ63134	Human recombinant
c 19	85	19.2	2030	17	AAT18996	Human steroid rece
c 20	85	19.2	2030	17	AAT30031	NER receptor poten
c 21	84.5	19.1	2958	20	AAX76729	Murine DIP coding
c 22	81	18.3	858	21	AAF18164	Lung cancer associ
c 23	80.5	18.2	65140	22	AAD17184	Streptomyces nours
c 24	80.5	18.2	125401	22	AAD17186	Streptomyces nours
c 25	79.5	18.0	1365	23	ABL23995	Drosophila melanog
26	79.5	18.0	2206	21	AAC74372	Human secreted pro
27	79.5	18.0	3549	23	ABL23994	Drosophila melanog
28	79	17.9	433	14	AAQ59506	Human brain Expres
c 29	79	17.9	1735	18	AAT77840	Human melanoma ass
c 30	79	17.9	2192	18	AAT77838	Human melanoma ass
31	79	17.9	5535	20	AAZ32057	Human METH2 relate
32	79	17.9	5535	20	AAX21355	Human BAI1 gene.
33	79	17.9	5535	22	AAC90314	AB005297 cDNA clon
c 34	78	17.6	1389	23	AAS84522	DNA encoding novel
c 35	78	17.6	1813	16	AAQ88760	Human ubiquitous n
c 36	78	17.6	1898	18	AAT79634	DNA encoding human
37	78	17.6	2400	23	AAS92478	DNA encoding novel
c 38	77.5	17.5	7725	22	AAK87157	Human immune/haema
c 39	77	17.4	3015	24	ABI99621	Mouse ischaemic co
40	77	17.4	17803	22	AAK68676	Human immune/haema
41	77	17.4	34269	22	AAK68677	Human immune/haema
42	77	17.4	34269	22	AAK85168	Human immune/haema
43	76.5	17.3	406	22	AAH42816	Nucleotide sequenc
44	76.5	17.3	1539	24	ABN95921	Gene #2419 used to
45	76.5	17.3	1539	24	ABL63095	Breast cancer rela

ALIGNMENTS

RESULT 1
ABL37082/c
ID ABL37082 standard; cDNA; 382 BP.
XX
AC ABL37082;
XX
DT 08-APR-2002 (first entry)
XX
DE Human colon tumour antigen polynucleotide SEQ ID NO:671.
XX
KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200196388-A2.
XX
PD 20-DEC-2001.

XX
PF 08-JUN-2001; 2001WO-US18557.
XX
PR 09-JUN-2000; 2000US-210899P.
PR 20-FEB-2001; 2001US-270216P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Harlocker SL, Secrist H;
XX
DR WPI; 2002-114514/15.
XX
PT Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient -
XX
PS Claim 1; SEQ ID 671; 105pp; English.
XX
CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC production. (I) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (I) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (I)
CC can be used in the diagnosis of a colon tumour.
XX
SQ Sequence 382 BP; 76 A; 94 C; 122 G; 86 T; 4 other;

Alignment Scores:

Pred. No.:	2.26e-38	Length:	382
Score:	442.00	Matches:	81
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-724-000A-5 (1-81) x ABL37082 (1-382)

```
Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 242

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 182

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 122

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 62

Qy      81 Val 81
           |||
Db      61 GTG 59
```

RESULT 2
ABK27569
ID ABK27569 standard; cDNA; 485 BP.
XX
AC ABK27569;
XX
DT 09-APR-2002 (first entry)
XX
DE Human colon cancer expressed sequence tag, Seq ID no 6.
XX
KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
KW expressed sequence tag.
XX
OS Homo sapiens.
XX
PN WO200196390-A2.
XX
PD 20-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US18577.
XX
PR 09-JUN-2000; 2000US-210821P.
PR 18-DEC-2000; 2000US-256571P.
PR 10-MAY-2001; 2001US-290240P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;
XX
DR WPI; 2002-139708/18.
XX
PT Novel isolated polynucleotide encoding a polypeptide comprising a
PT portion of colon tumour protein, useful for detection, diagnosis and
PT therapy of human colon cancer -
XX
PS Claim 1; Page 151-152; 220pp; English.
XX
CC The invention relates to an isolated polynucleotide (I) encoding a
CC polypeptide (II) comprising at least a portion of a colon tumour
CC protein. (I), (II) and antibody (III) to (II) are useful for determining
CC the presence of a cancer in a patient. (I), (II) or antigen presenting
CC cells expressing (I) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein, by contacting T cells with (I), (II) or
CC antigen-presenting cells that express (I), under conditions and for a
CC time sufficient to permit the stimulation and/or expansion of T cells.
CC (I), (II), or antigen presenting cells that express (II) are useful for
CC treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
CC isolated from a patient with (I), (II) or antigen presenting cells that
CC express (II), such that T cells proliferate, and administering to the
CC patient an effective amount of the proliferated T cells, thus inhibiting
CC the development of a cancer in the patient. (I) or (II) is useful in
CC vaccines and pharmaceutical compositions for prevention and treatment
CC of colon malignancies and for the diagnosis and monitoring of such
CC cancers. (I), (II) or (III) is useful for detection, diagnosis and/or
CC therapy of human colon cancer. (I) is useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of (II) in tumour cells.
CC ABK27564-ABK27807 represent novel human colon cancer coding
CC sequences and primers of the invention.
XX
SQ Sequence 485 BP; 109 A; 155 C; 124 G; 97 T; 0 other;

Alignment Scores:

Pred. No.:	3.07e-38	Length:	485
Score:	442.00	Matches:	81
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-724-000A-5 (1-81) x ABK27569 (1-485)

```
Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      79 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 138

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     139 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 198

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     199 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 258

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    259 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 318

Qy     81 Val 81
          |||
Db    319 GTG 321

RESULT 3
AAS21318
ID   AAS21318 standard; cDNA; 804 BP.
XX
AC   AAS21318;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human cDNA sequence encoding for PRO3446 polypeptide.
XX
KW   Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW   breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW   cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW   adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200140466-A2.
XX
PD   07-JUN-2001.
XX
PF   01-DEC-2000; 2000WO-US32678.
XX
PR   01-DEC-1999;   99WO-US28301.
PR   01-DEC-1999;   99WO-US28634.
PR   02-DEC-1999;   99WO-US28551.
PR   02-DEC-1999;   99WO-US28564.
PR   02-DEC-1999;   99WO-US28565.
PR   09-DEC-1999;   99US-0170262.
PR   16-DEC-1999;   99WO-US30095.
PR   20-DEC-1999;   99WO-US30911.
PR   20-DEC-1999;   99WO-US30999.
PR   30-DEC-1999;   99WO-US31243.
PR   06-JAN-2000; 2000WO-US00277.
PR   06-JAN-2000; 2000WO-US00376.
PR   11-FEB-2000; 2000WO-US03565.
PR   18-FEB-2000; 2000WO-US04341.
PR   18-FEB-2000; 2000WO-US04342.
PR   22-FEB-2000; 2000WO-US04414.
PR   24-FEB-2000; 2000WO-US04914.
PR   24-FEB-2000; 2000WO-US05004.
PR   01-MAR-2000; 2000WO-US05601.
PR   20-MAR-2000; 2000WO-US07377.
PR   21-MAR-2000; 2000WO-US07532.
PR   30-MAR-2000; 2000WO-US08439.
PR   17-MAY-2000; 2000WO-US13705.
PR   22-MAY-2000; 2000WO-US14042.
PR   30-MAY-2000; 2000WO-US14941.
PR   02-JUN-2000; 2000WO-US15264.
PR   10-NOV-2000; 2000WO-US30873.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Beresini M,  Deforge L,  Desnoyers L,  Filvaroff E,  Gao W;
PI   Gerritsen ME,  Goddard A,  Godowski PJ,  Gurney AL,  Sherwood S;
PI   Smith V,  Stewart TA,  Tumas D,  Watanabe CK,  Wood WI,  Zhang Z;
XX
DR   WPI; 2001-408281/43.
DR   P-PSDB; AAU12246.
XX
PT   Isolated, secretory and transmembrane PRO polypeptide used to detect
PT   other PRO polypeptides, link bioactive molecules to cells expressing
PT   PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT   lung, breast, prostate, cervical  -
XX
PS   Claim 3; Fig 149; 813pp; English.
XX
CC   AAS21244-AAS21518 encode for novel human secretory and transmembrane
CC   PRO polypeptides. The PRO polypeptides are useful to detect other
CC   PRO polypeptides, to link bioactive molecules to cells expressing
CC   PRO polypeptides, to modulate biological activities of cells expressing
CC   PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC   breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC   polypeptide expression in a cell sample to that in a control sample.
CC   Some of the 275 sequences are also useful to stimulate the release of
CC   tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC   proliferation or differentiation of chondrocytes, the proliferation or
CC   gene expression in pericyte cells, the release of proteoglycans from
CC   cartilage, the proliferation of inner ear utricular supporting cells or
CC   of T-lymphocytes, the release of a cytokine from peripheral blood
CC   monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC   the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC   skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC   to factor VIIA. The PRO polypeptides can be used in assays to identify
CC   molecules involved in binding interactions. The polynucleotides encoding
CC   PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC   transgenic or knock out animals and can be used in gene therapy.
XX
SQ   Sequence 804 BP; 199 A; 263 C; 176 G; 166 T; 0 other;

Alignment Scores:
Pred. No.:              5.86e-38      Length:          804
Score:                  442.00        Matches:         81
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     22            Gaps:            0

US-09-724-000A-5 (1-81) x AAS21318 (1-804)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy     81 Val 81
          |||
Db    282 GTG 284

RESULT 4
AAD27025
ID   AAD27025 standard; DNA; 806 BP.
XX
AC   AAD27025;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human Secreted epithelial colon stromal-1 (Secs-1) DNA.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   human; ds.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             29..274
FT                   /*tag=  a
FT                   /product= "Human Secs-1 protein"
FT   sig_peptide     29..100
FT                   /*tag=  b
FT   mat_peptide     101..271
FT                   /*tag=  c
FT                   /product= "Mature human Secs-1 protein"
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
DR   P-PSDB; AAE16481.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Claim 1; Fig 2; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is human Secs-1
CC   DNA.
XX
SQ   Sequence 806 BP; 207 A; 257 C; 179 G; 163 T; 0 other;

Alignment Scores:
Pred. No.:                5.88e-38        Length:         806
Score:                    442.00          Matches:        81
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       24              Gaps:           0

US-09-724-000A-5 (1-81) x AAD27025 (1-806)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 88

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 148

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     149 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 208

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     209 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 268

Qy      81 Val 81
           |||
Db     269 GTG 271

RESULT 5
AAF97898
ID   AAF97898 standard; cDNA; 908 BP.
XX
AC   AAF97898;
XX
DT   01-JUN-2001  (first entry)
XX
DE   Human secreted protein cDNA, SEQ ID NO: 25.
XX
KW   Human; secreted protein; immunomodulatory; antisclerotic;
KW   dermatological; antiinflammatory; anti-HIV; cytostatic; cardiant;
KW   vascular; anti-angiogenic; ophthalmological; neuroprotectant;
KW   nootropic; anticonvulsant; antialzheimers; antiparkinsonian;
KW   antimicrobial; vulnerary; vaccine; gene therapy; cancer;
KW   protein coordinate data; infection; ss.
XX
OS   Homo sapiens.
XX
PN   WO200121658-A1.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-US26013.
XX
PR   24-SEP-1999;   99US-0155709.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J,  Baker KP,  Birse CE,  Ebner R,  Fiscella M,  Komatsoulis GA;
PI   Lafleur DW,  Moore PA,  Olsen HS,  Rosen CA,  Ruben SA,  Soppet DR;
PI   Young PE,  Wei P,  Florence KA;
XX
DR   WPI; 2001-235311/24.
XX
PT   Nucleic acids encoding 32 human secreted polypeptides, useful for
PT   preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT   disease and diabetic retinopathy -
XX
PS   Claim 1; Page 730; 890pp; English.
XX
CC   The present sequence encodes one of 32 novel human secreted polypeptides.
CC   The nucleic acid molecules and polypeptides they encode may be used in
CC   the prevention, diagnosis and treatment of diseases such as
CC   immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus
CC   and human immuno-deficiency virus (HIV) infections), hyperproliferative
CC   disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
CC   (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary
CC   arteriosclerosis), angiogenic disorders (e.g. corneal graft
CC   neovascularisation and diabetic retinopathy), neurological disorders
CC   (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
CC   infectious diseases and/or for promoting wound healing, regeneration
CC   and/or chemotaxis. The nucleic acid molecules may be used to produce the
CC   secreted polypeptides. They may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acid
CC   sequences in samples. The polypeptides may be used as antigens in the
CC   production of antibodies and in assays to identify modulators of
CC   their expression and activity.
XX
SQ   Sequence 908 BP; 242 A; 290 C; 197 G; 175 T; 4 other;

Alignment Scores:
Pred. No.:                6.85e-38        Length:         908
Score:                    442.00          Matches:        81
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
```

```
Query Match:            100.00%       Indels:        0
DB:                     22            Gaps:          0

US-09-724-000A-5 (1-81) x AAF97898 (1-908)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 144

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 204

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 264

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 324

Qy     81 Val 81
          |||
Db    325 GTG 327

RESULT 6
AAA72230/c
ID   AAA72230 standard; cDNA; 797 BP.
XX
AC   AAA72230;
XX
DT   06-DEC-2000  (first entry)
XX
DE   Human CASB gene partial cDNA, SEQ ID NO:7.
XX
KW   Human; CASB gene; overexpression; colon tumour-associated antigen;
KW   expressed sequence tag; EST; colon cancer; tumour; autoimmune disease;
KW   diagnosis; disease susceptilibity; prophylaxis; genetic vaccine;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043509-A2.
XX
PD   27-JUL-2000.
XX
PF   17-JAN-2000; 2000WO-EP00346.
XX
PR   19-JAN-1999;   99GB-0001078.
PR   29-JAN-1999;   99GB-0002090.
PR   01-FEB-1999;   99GB-0002163.
PR   01-FEB-1999;   99GB-0002168.
PR   01-FEB-1999;   99GB-0002169.
PR   07-APR-1999;   99GB-0007901.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Vinals-Bassols C;
XX
DR   WPI; 2000-482912/42.
XX
PT   New isolated polynucleotide useful for diagnosis and/or treatment of
PT   colon cancer and autoimmune disease -
XX
PS   Disclosure; Page 35-36; 41pp; English.
XX
CC   Sequences AAA72230-A72235 represent human CASB gene partial cDNA
CC   sequences which are derived from expressed sequence tags (ESTs). The
CC   invention relates to human CASB cDNA sequences CASB611, CASB500,
CC   CASB501, CASB502, CASB505 and CASB507 (AAA72224-A72229, respectively) and
CC   also to these human CASB partial cDNA sequences. Expression of the human
CC   CASB genes (with the exception of CASB611) is associated with colon
CC   tumours, and the encoded proteins (sequences not given in the
CC   specification) represent colon tumour-associated antigens. The cDNA
CC   sequences may be used in diagnosing the presence or a susceptibility to
CC   a disease related to the presence, expression or activity of CASB genes.
CC   Such diseases include autoimmune diseases and especially colon cancer.
CC   The nucleic acid sequences may also be used in genetic vaccines for the
CC   prophylaxis or therapeutic treatment of colon cancer and autoimmune
CC   diseases.
XX
SQ   Sequence 797 BP; 164 A; 180 C; 254 G; 199 T; 0 other;

Alignment Scores:
Pred. No.:              1.77e-33      Length:        797
Score:                  400.00        Matches:       78
Percent Similarity:     96.34%        Conservative:  1
Best Local Similarity:  95.12%        Mismatches:    2
Query Match:            90.50%        Indels:        2
DB:                     21            Gaps:          0

US-09-724-000A-5 (1-81) x AAA72230 (1-797)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    765 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 706

Qy     21 SerThrGluGlyLysArgArgProAlaLys-AlaTrpSerGlyArgArgThrArgLeuCy 40
          ||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||
Db    705 TCCACAGAGGGGAAGAGGCGTCCTGCCAACGGCCTGGTCAGGCAGGAGAACCAGGCTCTG 646

Qy     40 sCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCy 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    645 CTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTG 586

Qy     60 sLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGl 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||   :::|||||
Db    585 TAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGA-GTCCCACA 527

Qy     80 nVal 81
          ||||
Db    526 GGTG 523

RESULT 7
AAA72224/c
ID   AAA72224 standard; cDNA; 801 BP.
XX
AC   AAA72224;
XX
DT   06-DEC-2000  (first entry)
XX
DE   Human CASB611 cDNA.
XX
KW   Human; CASB611; colon-specific expression; expressed sequence tag;
KW   EST; colon cancer; tumour; autoimmune disease; diagnosis;
KW   disease susceptilibity; prophylaxis; genetic vaccine; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043509-A2.
XX
PD   27-JUL-2000.
XX
PF   17-JAN-2000; 2000WO-EP00346.
XX
PR   19-JAN-1999;   99GB-0001078.
PR   29-JAN-1999;   99GB-0002090.
PR   01-FEB-1999;   99GB-0002163.
PR   01-FEB-1999;   99GB-0002168.
PR   01-FEB-1999;   99GB-0002169.
PR   07-APR-1999;   99GB-0007901.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Vinals-Bassols C;
XX
```

```
DR   WPI; 2000-482912/42.
XX
PT   New isolated polynucleotide useful for diagnosis and/or treatment of
PT   colon cancer and autoimmune disease -
XX
PS   Claim 3; Page 34; 41pp; English.
XX
CC   This sequence represents human CASB611 cDNA. This gene exhibits
CC   colon-specific expression and is highly expressed in the rectum.
CC   The invention relates to human CASB cDNA sequences CASB611, CASB500,
CC   CASB501, CASB502, CASB505 and CASB507 (AAA72224-A72229, respectively)
CC   and also to human CASB partial cDNA sequences (AAA72230-A72235) derived
CC   from expressed sequence tags (ESTs). Expression of the human CASB genes
CC   (with the exception of CASB611) is associated with colon tumours, and
CC   the encoded proteins (sequences not given in the specification)
CC   represent colon tumour-associated antigens. The cDNA sequences may be
CC   used in diagnosing the presence or a susceptibility to a disease related
CC   to the presence, expression or activity of CASB genes. Such diseases
CC   include autoimmune diseases and especially colon cancer. The nucleic
CC   acid sequences may also be used in genetic vaccines for the prophylaxis
CC   or therapeutic treatment of colon cancer and autoimmune diseases.
XX
SQ   Sequence 801 BP; 165 A; 181 C; 255 G; 200 T; 0 other;

Alignment Scores:
Pred. No.:              1.79e-33        Length:         801
Score:                  400.00          Matches:        78
Percent Similarity:     96.34%          Conservative:   1
Best Local Similarity:  95.12%          Mismatches:     2
Query Match:            90.50%          Indels:         2
DB:                     21              Gaps:           0

US-09-724-000A-5 (1-81) x AAA72224 (1-801)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 710

Qy      21 SerThrGluGlyLysArgArgProAlaLys-AlaTrpSerGlyArgArgThrArgLeuCy 40
           ||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||
Db     709 TCCACAGAGGGGAAGAGGCGTCCTGCCAACGGCCTGGTCAGGCAGGAGAACCAGGCTCTG 650

Qy      40 sCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCy 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 CTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTG 590

Qy      60 sLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGl 80
           ||||||||||||||||||||||||||||||||||||||||||||||||   :::|||||
Db     589 TAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGA-GTCCCACA 531

Qy      80 nVal 81
           ||||
Db     530 GGTG 527

RESULT 8
AAD27024
ID   AAD27024 standard; DNA; 744 BP.
XX
AC   AAD27024;
XX
DT   09-APR-2002  (first entry)
XX
DE   Mouse Secreted epithelial colon stromal-1 (Secs-1) gene.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   mouse; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             38..274
FT                   /*tag=  a
FT                   /product= "Mouse Secs-1protein"
FT   sig_peptide     38..109
FT                   /*tag=  b
FT   mat_peptide     110..271
FT                   /*tag=  c
FT                   /product= "Mature mouse Secs-1 protein"
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
DR   P-PSDB; AAE16479.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Claim 1; Fig 1; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is mouse Secs-1
CC   gene.
XX
SQ   Sequence 744 BP; 209 A; 222 C; 155 G; 158 T; 0 other;

Alignment Scores:
Pred. No.:              3.37e-15        Length:         744
Score:                  228.50          Matches:        45
Percent Similarity:     67.90%          Conservative:   10
Best Local Similarity:  55.56%          Mismatches:     23
Query Match:            51.70%          Indels:         3
DB:                     24              Gaps:           1

US-09-724-000A-5 (1-81) x AAD27024 (1-744)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           |||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db      38 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 97

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           |||:::||||||:::|||   |||||||||:::         ||||||         |||
Db      98 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 148

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||      |||            |||      ||||||:::|||   |||   |||
Db     149 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 208
```

```
Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           :::   |||:::   :::       :::   ||||||||||||||||||||||||
Db     209 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 268

Qy      81 Val 81
           :::
Db     269 ATA 271

RESULT 9
AAD27026
ID   AAD27026 standard; DNA; 4159 BP.
XX
AC   AAD27026;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human Secreted epithelial colon stromal-1 (Secs-1) gene.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4159
FT                   /*tag=  a
FT                   /product= "Human Secs-1 protein"
FT                   /note= "CDS does not include stop codon"
FT                   /partial
FT   exon            1..69
FT                   /*tag=  d
FT                   /note= "Encodes Secs-1 peptide (AAE16484)"
FT   intron          70..2626
FT                   /*tag=  d
FT                   /cons_splice= (5'site:NO, 3'site:NO)
FT   exon            2627..2725
FT                   /*tag=  d
FT                   /note= "Encodes Secs-1 peptide (AAE16485)"
FT   intron          2726..4075
FT                   /*tag=  d
FT                   /cons_splice= (5'site:NO, 3'site:NO)
FT   exon            4076..4159
FT                   /*tag=  d
FT                   /note= "Encodes Secs-1 peptide (AAE16486)"
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
DR   P-PSDB; AAE16481, AAE16484, AAE16485, AAE16486.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Disclosure; Fig 4; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is human Secs-1
CC   gene.
XX
SQ   Sequence 4159 BP; 1009 A; 1037 C; 983 G; 1110 T; 20 other;

Alignment Scores:
Pred. No.:                7.58e-09      Length:          4159
Score:                    178.00        Matches:         37
Percent Similarity:       62.12%        Conservative:    4
Best Local Similarity:    56.06%        Mismatches:      13
Query Match:              40.27%        Indels:          13
DB:                       24            Gaps:            1

US-09-724-000A-5 (1-81) x AAD27026 (1-4159)

Qy      24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArg 43
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2635 GGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGA 2694

Qy      44 ValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCysLysProCys 63
           ||||||||||||||||||||||||||||||||.  :::      |||   :::   |||
Db    2695 GTCCCTAGCCCCAACTCAACAAACCTGAAAGGTAAGTACCCCCACCTCGTCCAGAC-TGT 2753

Qy      64 LysLeuGluPro----------------------------------GluProArgLeu 71
              |||                                        |||
Db    2754 GGGGCAGAAGTTCTACAGTGGCCATGGGACCAGCCACACACACTGATCAGCCCCCACCCA 2813

Qy      72 TrpValValProGlyAla 77
           |||:::      |||:::
Db    2814 TGGCTGGCATCAGGCTCT 2831

RESULT 10
AAS39569
ID   AAS39569 standard; cDNA; 487 BP.
XX
AC   AAS39569;
XX
DT   17-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human colon associated polypeptide #222.
XX
KW   Human; colon cancer; congenital abnormality; infection; colitis;
KW   inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KW   intestinal inflammatory disorder; malabsorption syndrome; gastric;
KW   sigmoid disease; antibacterial; antiviral; antiinflammatory;
KW   cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155302-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01240.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
```

PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465567/50.
DR P-PSDB; AAU22689.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the colon including colon cancers and also for
PT testing and detection e.g. diagnosis -
XX
PS Claim 4; SEQ ID No 232; 562pp; English.
XX
CC The present invention relates to the isolation of novel human colon
CC associated polypeptides (AAU22468-AAU22701), and the cDNA and genomic

CC sequences encoding for them. The sequences of the invention are useful
CC in the diagnosis, treatment, prevention and/or prognosis of disorders
CC of the colon including colon cancer, congenital abnormalities
CC (e.g. atresia and stenosis), bacterial and viral infections,
CC inflammatory bowel disease (IBD), neoplastic cell disorders
CC (e.g. polyps and adenomas, intestinal inflammatory disorders, colitis,
CC colonic inflammation, diarrhoea and dysentery, malabsorption syndromes
CC (e.g. lactose intolerance), intestinal obstruction and sigmoid diseases.
CC The polynucleotides sequences of the invention can also be used in gene
CC therapy. AAS39348-AAS39581 represent cDNA sequences encoding for the
CC novel human colon associated polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 487 BP; 135 A; 156 C; 100 G; 95 T; 1 other;

Alignment Scores:

Pred. No.:	0.741	Length:	487
Score:	92.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	20.81%	Indels:	0
DB:	22	Gaps:	0

US-09-724-000A-5 (1-81) x AAS39569 (1-487)

```
Qy     38 ArgLeuCysCysHisArgValProSerProAsnSerThrAsnLeuLys 53
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 AGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAA 50
```

RESULT 11
AAK88908
ID AAK88908 standard; cDNA; 487 BP.
XX
AC AAK88908;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen coding sequence SEQ ID NO: 1224.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ss.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-502630/55.
DR P-PSDB; AAM93135.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognozing disorders of the
PT digestive system, particularly cancer and cancer metastases -
XX
PS Claim 1; SEQ ID NO 1224; 986pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a cDNA encoding a digestive
CC system antigen of the invention.
XX
SQ Sequence 487 BP; 135 A; 156 C; 100 G; 95 T; 1 other;

Alignment Scores:

Pred. No.:	0.741	Length:	487
Score:	92.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	20.81%	Indels:	0

DB: 22 Gaps: 0

US-09-724-000A-5 (1-81) x AAK88908 (1-487)

```
Qy     38 ArgLeuCysCysHisArgValProSerProAsnSerThrAsnLeuLys 53
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 AGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAA 50
```

RESULT 12
ABK41859/c
ID ABK41859 standard; cDNA; 619 BP.
XX
AC ABK41859;
XX
DT 21-MAY-2002 (first entry)
XX
DE cDNA encoding novel human connective tissue related polypeptide #247.
XX
KW Human; connective tissue related disorder; cancer; gene therapy;
KW cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200155343-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01322.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-565190/63.
DR P-PSDB; AAU86681.
XX
PT Nucleic acid encoding novel connective tissue associated polypeptides,
PT used in diagnosing, preventing, treating or ameliorating a disorder
PT such as cancer or rheumatoid arthritis -
XX
PS Claim 4; SEQ ID No 257; 673pp; English.
XX
CC The present invention relates to the isolation of novel human connective
CC tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
CC (cDNA and genomic) sequences encoding them. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of diseases associated with connective tissue(s), including
CC cancer. The polynucleotide sequences of the invention are also useful
CC in gene therapy. ABK41613-ABK42101 represent cDNA sequences encoding
CC the novel human connective tissue related polypeptides.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 619 BP; 129 A; 194 C; 191 G; 103 T; 2 other;

Alignment Scores:

Pred. No.:	1.29	Length:	619
Score:	91.00	Matches:	24
Percent Similarity:	37.97%	Conservative:	6
Best Local Similarity:	30.38%	Mismatches:	23
Query Match:	20.59%	Indels:	26
DB:	23	Gaps:	3

US-09-724-000A-5 (1-81) x ABK41859 (1-619)

```
Qy       3 LeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePheSerThr 22
           ||||||   |||   ||||||   |||:::|||   ||||||   :::
Db     426 CTTCTGGCTCTGCAGTCGCTTCAGTGTCTCCTCCACCTCTGCCTCACAGAGCTGGGAGGA 367

Qy      23 GluGlyLysArg-----------------------------------------ArgPro 28
             |||                                                 |||
Db     366 GGCGGAGGAAATACAAGATACATAAAGTGCTTGGCAGAGCCTGCCACAAACTGAGCTCCC 307

Qy      29 AlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
           |||   |||   |||                 |||||||||   |||   |||
Db     306 GGGAAACGTTGGCAGGGT-----------------TGCCATCGCGCACCAAGGCCCCAA 265

Qy      49 SerThrAsnLeuLysGly-----------------HisHisValArgLeuCysLys 61
           :::      |||:::|||                 ||||||:::   |||   :::
Db     264 AACWCTAGGCTGCAGGGCCGCCGCCCACCCCTGGCCCATCATCTCTTCCTGGCCGAG 208
```

```
RESULT 13
AAQ57015
ID   AAQ57015 standard; DNA; 2167 BP.
XX
AC   AAQ57015;
XX
DT   31-AUG-1994  (first entry)
XX
DE   PKC gamma.
XX
KW   110 kD catalytic subunit; phosphatidyl inositol 3-kinase;
KW   transformation; Schizosaccharomyces pombe; nmt promoter; thiamine;
KW   PtdIns 3-kinase; assay; detection; cell growth; regulation; cancer;
KW   blood vessel plaques; ss.
XX
OS   Bos taurus.
XX
PN   WO9403609-A.
XX
PD   17-FEB-1994.
XX
PF   05-AUG-1993;   93WO-GB01651.
XX
PR   05-AUG-1992;   92GB-0016654.
XX
PA   (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
XX
PI   Goode NT,  Nurse PM,  Parker PJJ,  Waterfield MD;
XX
DR   WPI; 1994-065697/08.
XX
PT   Eukaryotic cells transformed with mammalian phospholipid or
PT   protein kinase DNA - useful in assays for compounds involved in
PT   cell growth regulation and for treating cancers
XX
PS   Disclosure; Page 41-42; 71pp; English.
XX
CC   The sequences given in AAQ57014-17 encode protein kinase C (PKC)
CC   epsilon, gamma, delta and nu respectively.  These sequences were
CC   transformed into Schizosaccharomyces pombe cells under the regulatory
CC   control of the nmt promoter in an embodiment of the invention.  In the
CC   presence of thiamine the promoter is inactive and the cells carrying
CC   the PKC plasmids grow as the parental strain.  In the absence
CC   of thiamine the nmt promoter functions and the PKC is induced.  PKC
CC   activity is substantially increased under these conditions.  Cells
CC   containing constructs such as this, are useful in assays for detecting
CC   compounds involved in cell growth regulation.  It is also used as the
CC   basis for detecting compounds for treating cancers and the formation
CC   of blood vessel plaques.
XX
SQ   Sequence 2167 BP; 428 A; 700 C; 611 G; 428 T; 0 other;

Alignment Scores:
Pred. No.:                19.5          Length:        2167
Score:                    86.50         Matches:       26
Percent Similarity:       39.76%        Conservative:  7
Best Local Similarity:    31.33%        Mismatches:    19
Query Match:              19.57%        Indels:        31
DB:                       15            Gaps:          4

US-09-724-000A-5 (1-81) x AAQ57015 (1-2167)

Qy     18 SerIlePheSerThrGluGlyLysArgArg---------ProAlaLysAlaTrp------ 32
          |||      |||:::   |||:::||||||         |||      ||||||
Db   1882 TCGACAAGTTCTTCACTCGGGCGGCGCCGGCGCTGACACCCCCTGACCGCCTGGTTCTGG 1941

Qy     33 ----------------------------SerGlyArgArgThrArgLeuCysCysHis 42
                                       |||       ||||||:::   |||
Db   1942 CCAGCATCGACCAGGCTGAGTTCCAGGGCTTCACCTATGTCAACCCGGATTTCGTGCACC 2001

Qy     43 ArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCysLysPro 62
          |||:::|||:::      |||                    |||||||||:::
Db   2002 CGGATGCCCGCAGCCCCATCAGCCCAA----------------CGCCTGTGCCAGTCA 2043

Qy     63 Cys----------------------------LysLeuGluProGluProArgLeuTrp 72
          |||                           :::|||   |||   ||||||||||||
Db   2044 TGTAATCCCACCTGCCGCCACCAGGCGTCCCCACGGCTCCCTCCTCCGCCCCGGCTTTGG 2103

Qy     73 ValValPro 75
          |||
Db   2104 CCCTCGCCT 2112

RESULT 14
AAC98948/c
ID   AAC98948 standard; cDNA; 1622 BP.
XX
AC   AAC98948;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:176.
XX
KW   Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW   detection; diagnosis; identification; cytostatic; neuroprotective;
KW   nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW   antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW   linkage analysis; tissue identification; tissue typing; forensic;
KW   neural; immune system; muscular; reproductive; gastrointestinal;
KW   pulmonary; cardiovascular; renal; proliferative; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055320-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05989.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-579444/54.
DR   P-PSDB; AAB54183.
XX
PT   New nucleic acid that is a pancreatic cancer antigen for preventing,
PT   treating, or ameliorating a medical condition, particular pancreatic
PT   cancer, or for use in assays for diagnosing a pathological condition -
XX
PS   Claim 1; Page 629; 1379pp; English.
XX
CC   AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC   proteins, called pancreatic cancer antigens, given in AAB54008 to
CC   AAB54466. The human pancreatic cancer antigens have cytostatic,
CC   neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC   gynaecological, cardiant and antiinflammatory activities, and can be used
CC   in gene therapy. The polynucleotide and proteins can be used for
CC   preventing, treating, or ameliorating a medical condition or in assays
CC   for diagnosing a pathological condition or a susceptibility to one in a
CC   subject. Binding partners to the proteins and the activity of the
CC   proteins can be identified. The pancreatic cancer antigens can be used to
CC   detect, treat or prevent pancreatic disorders, especially cancer.
CC   Agonists and antagonists to the antigens can be screened for. The
CC   pancreatic cancer antigen polynucleotides can be used to design nucleic
CC   acid hybridisation probes that can be used in chromosome mapping, linkage
CC   analysis, tissue identification and/or typing and a variety of forensic
CC   and diagnostic methods. The proteins can be used to generate antibodies
CC   which are used to purify, detect and target the polypeptides, including
CC   both in vivo and in vitro diagnostic and therapeutic methods. The
CC   proteins can be used to treat or prevent neural, immune system, muscular,
CC   reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC   proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC   sequences used in the exemplification of the present invention.
XX
```

```
SQ   Sequence 1622 BP; 328 A; 486 C; 515 G; 273 T; 20 other;

Alignment Scores:
Pred. No.:              19.4          Length:        1622
Score:                  85.00         Matches:       26
Percent Similarity:     47.76%        Conservative:  6
Best Local Similarity:  38.81%        Mismatches:    17
Query Match:            19.23%        Indels:        18
DB:                     21            Gaps:          4

US-09-724-000A-5 (1-81) x AAC98948 (1-1622)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||       |||:::|||      |||:::|||||||||      :::
Db    1022 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 963

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||      |||   |||   |||   :::   :::|||
Db     962 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 921

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                        |||||||||                |||||||||
Db     920 --------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCGCCTTCCCCGGAGC 876

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     875 CCTGGCTGCCTGCCTCAGATC 855

RESULT 15
AAT27617/c
ID   AAT27617 standard; cDNA to mRNA; 1688 BP.
XX
AC   AAT27617;
XX
DT   06-NOV-1996  (first entry)
XX
DE   Steroid hormone receptor analogue ECDN small mol. variant cDNA.
XX
KW   Human; foetal lung; steroid hormone; receptor; analogue protein;
KW   ECDN protein; cancer; screening; binding molecule; recombinant;
KW   identification; anticancer drug; cancerous tissue; primer;
KW   probe; antibody; immunohistochemical assay; variant;
KW   small molecule; ECDNsm; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             206..1300
FT                   /*tag=  a
XX
PN   WO9609324-A1.
XX
PD   28-MAR-1996.
XX
PF   21-SEP-1995;   95WO-JP01909.
XX
PR   21-SEP-1994;   94JP-0226270.
XX
PA   (EISA ) EISAI CO LTD.
PA   (GANK-) ZH GAN KENKYUKAI.
PA   (CANC-) CANCER INST.
XX
PI   Nakamura Y,  Saito H;
XX
DR   WPI; 1996-188403/19.
DR   P-PSDB; AAR96235.
XX
PT   ECDN protein, a steroid hormone receptor analogue from human foetal
PT   lung - is expressed in cancer cells and is useful for cancer
PT   diagnosis and drug development
XX
PS   Claim 3; Pages 25-28; 43pp; Japanese.
XX
CC   The present sequence encodes the variant of the human foetal lung
CC   derived steroid hormone receptor analogue protein ECDN, designated
CC   ECDN small mol. (ECDNsm) protein. ECDNsm protein is expressed in
CC   various cancer cells, therefore screening for ECDNsm protein
CC   binding mols., using recombinant ECDNsm proteins will be useful in
CC   the identification of candidate anticancer drugs. Gene expression
CC   of ECDNsm proteins in cancerous tissues can be studied using
CC   primers and probes derived from ECDNsm protein cDNA. Antibodies
CC   which recognise ECDNsm proteins can be used in ECDNsm protein
CC   immunohistochemical assays.
XX
SQ   Sequence 1688 BP; 347 A; 562 C; 494 G; 285 T; 0 other;

Alignment Scores:
Pred. No.:              20.4          Length:        1688
Score:                  85.00         Matches:       26
Percent Similarity:     47.76%        Conservative:  6
Best Local Similarity:  38.81%        Mismatches:    17
Query Match:            19.23%        Indels:        18
DB:                     17            Gaps:          4

US-09-724-000A-5 (1-81) x AAT27617 (1-1688)

Qy      22 ThrGluGlyLysArgArgProAlaLysAla------TrpSerGlyArgArgThrArgLeu 39
           |||   |||   |||       |||:::|||      |||:::|||||||||      :::
Db     692 ACTGGGGGTCTGCGCCCAGGGGCCAGGGCGTGACTTTGGGCTGGTCGGAGAAGGAGCGTT 633

Qy      40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||      |||   |||   |||   :::   :::|||
Db     632 TGTTGCACTGCAGTTGGGCCGCCACCAACTGCTGGATCATTA----------------- 591

Qy      59 LeuCysLysProCysLysLeuGluPro-----------GluProArgLeuTrpValVal 74
                        |||||||||                |||||||||
Db     590 --------------GTTCTTGAGCCGCTGTTAGCTGGACACCCTCGCCTTCCCCGGAGC 546

Qy      75 ProGlyAlaLeuProGlnVal 81
           ||||||   |||||||||:::
Db     545 CCTGGCTGCCTGCCTCAGATC 525


Search completed: May 17, 2003, 11:17:01
Job time : 166 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 12:22:51 ; Search time 1740 Seconds
(without alignments)
7502.047 Million cell updates/sec

Title: US-09-724-000A-4
Perfect score: 806
Sequence: 1 ggaacgagggaaaatctgcc.....actcaatgcagacacaaaaa 806

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	486	60.3	557	14	BQ189412	BQ189412 UI-E-EJ1-
c	2	460	57.1	460	9	AI983767	AI983767 wu20a04.x
c	3	450	55.8	453	10	AW516596	AW516596 xq01h07.x
	4	429	53.2	906	12	BE899580	BE899580 601682443
c	5	410	50.9	543	10	AW970357	AW970357 EST382438
c	6	406	50.4	413	9	AI339648	AI339648 qk63a12.x
c	7	398	49.4	450	9	AI948903	AI948903 wq17c02.x
c	8	393	48.8	444	9	AI833297	AI833297 at67a07.x
c	9	380	47.1	467	9	AI304380	AI304380 qo59c12.x
c	10	379	47.0	430	9	AI283185	AI283185 qk49g09.x
c	11	377	46.8	429	9	AI813445	AI813445 wj06e01.x
c	12	371	46.0	517	10	AW854263	AW854263 RC3-CT025
c	13	369	45.8	420	9	AA938765	AA938765 on44h11.s
c	14	366	45.4	455	9	AI832391	AI832391 at68f03.x
c	15	362	44.9	467	9	AI831407	AI831407 wj64a12.x
c	16	359	44.5	405	9	AI833021	AI833021 at74c12.x
c	17	353	43.8	423	9	AI336470	AI336470 qo61d01.x
	18	351	43.5	466	10	AW361500	AW361500 QV2-CT026
c	19	350	43.4	416	10	AW206923	AW206923 UI-H-BI1-
c	20	346	42.9	431	9	AI832498	AI832498 at69e07.x
c	21	342	42.4	382	9	AI732377	AI732377 nf90e09.x
c	22	333	41.3	384	9	AI336626	AI336626 qo62d09.x
c	23	323	40.1	331	9	AI246768	AI246768 qk40f07.x
c	24	323	40.1	381	9	AI285352	AI285352 ql38g09.x
	25	321	39.8	356	10	AW351839	AW351839 RC1-CT019
c	26	320	39.7	381	9	AI732376	AI732376 nf80f04.x
c	27	313	38.8	330	10	AW351854	AW351854 RC1-CT019
c	28	308	38.2	476	10	AW134688	AW134688 UI-H-BI1-
c	29	307	38.1	460	12	BF001316	BF001316 7g62b10.x
c	30	306	38.0	400	9	AI864896	AI864896 wj66d04.x
c	31	300	37.2	427	10	AW361498	AW361498 QV2-CT026
c	32	288	35.7	441	9	AA587764	AA587764 nm95f05.s
	33	286	35.5	302	12	BF478262	BF478262 nae45a11.
	34	281	34.9	503	9	AA422178	AA422178 zv31g07.r
c	35	277	34.4	484	10	AW970275	AW970275 EST382356
c	36	261	32.4	394	9	AI744428	AI744428 wf89b03.x
c	37	260	32.3	292	10	AW361503	AW361503 QV2-CT026
c	38	259	32.1	357	9	AI766378	AI766378 wh61c01.x
c	39	249	30.9	487	9	AA553959	AA553959 nl01d03.s
c	40	246	30.5	429	9	AA857922	AA857922 of69c03.s
c	41	230	28.5	415	10	AW361502	AW361502 QV2-CT026
c	42	228	28.3	381	9	AA534511	AA534511 nf80f04.s
c	43	227	28.2	278	9	AA535314	AA535314 nf90e09.s
c	44	218	27.0	385	10	AW591238	AW591238 xo14a06.x
c	45	196	24.3	366	9	AA422086	AA422086 zv31g06.s

ALIGNMENTS

RESULT 1
BQ189412
LOCUS BQ189412 557 bp mRNA linear EST 30-APR-2002
DEFINITION UI-E-EJ1-ajv-h-16-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone UI-E-EJ1-ajv-h-16-0-UI 5', mRNA sequence.
ACCESSION BQ189412
VERSION BQ189412.1 GI:20364963
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 557)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

```
           Clone Distribution: Researchers may obtain clones from Research
           Genetics (www.resgen.com).
           Seq primer: M13 REVERSE.
FEATURES             Location/Qualifiers
     source          1. .557
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EJ1-ajv-h-16-0-UI"
                     /clone_lib="UI-E-EJ1"
                     /tissue_type="fetal eyes, lens, eye anterior segment,
                     optic nerve, retina, Retina Foveal and Macular, RPE and
                     Choroid"
                     /dev_stage="fetal and adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EJ1 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are: fetal eyes, AGAATCAAGA
                     ; lens, CGATTAGCGA; eye anterior segment, AATGCCGCAT;
                     optic nerve, CCATTAAGTG; retina, CCGCG; Retina Foveal and
                     Macular, GTCC; RPE and Choroid, ACCTA. This library was
                     created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
BASE COUNT      135 a    182 c    138 g    101 t      1 others
ORIGIN

  Query Match             60.3%;  Score 486;  DB 14;  Length 557;
  Best Local Similarity   99.8%;  Pred. No. 3.4e-232;
  Matches  536;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      83 ATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGG 142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGG 60

Qy     143 CTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGG 202
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGG 120

Qy     203 CTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTC 262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTC 180

Qy     263 CCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACCTCATGCCTGGCAC 322
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     181 CCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAAAACCTCATGCCTGGCAC 240

Qy     323 CTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATG 382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATG 300

Qy     383 TTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTT 442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTT 360

Qy     443 GTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCT 502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCT 420

Qy     503 CAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTC 562
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTC 480

Qy     563 ACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCA 619
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCA 537
```

```
RESULT 2
AI983767/c
LOCUS       AI983767                460 bp    mRNA    linear   EST 27-OCT-1999
DEFINITION  wu20a04.x1 Soares_Dieckgraefe_colon_NHCD Homo sapiens cDNA clone
            IMAGE:2520558 3', mRNA sequence.
ACCESSION   AI983767
VERSION     AI983767.1  GI:5810986
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2520558"
                     /clone_lib="Soares_Dieckgraefe_colon_NHCD"
                     /tissue_type="colonic mucosa from 3 patients with Crohn's
                     disease"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCGTCTTTTTTTTTTTTTTTTT 3'),
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Tissue samples
                     provided by Dr. Brian Dieckgraefe (Washington University,
                     dieck@im.wustl.edu); colonic mucosa represents a range of
                     disease involvement from moderate to severe Crohn's
                     disease; samples include both perforating (fistulas) and
                     non-perforating samples. Library constructed by Bento
                     Soares and M. Fatima Bonaldo.  "
BASE COUNT       96 a     94 c    149 g    121 t
ORIGIN

  Query Match             57.1%;  Score 460;  DB 9;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-219;
  Matches  460;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     331 CCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGG 390
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     460 CCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGG 401

Qy     391 CTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAAT 450
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     400 CTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAAT 341

Qy     451 TGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCC 510
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     340 TGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCC 281

Qy     511 ACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTAC 570
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     280 ACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTAC 221

Qy     571 AAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGA 630
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     220 AAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGA 161
```

```
Qy    631 CCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACC 690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 CCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACC 101

Qy    691 AGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAG 750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 AGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAG 41

Qy    751 AGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTC 790
          |||||||||||||||||||||||||||||||||||||||
Db     40 AGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTC 1
```

```
RESULT 3
AW516596/c
LOCUS       AW516596                453 bp    mRNA    linear   EST 03-MAR-2000
DEFINITION  xq01h07.x1 Soares_NHCeC_cervical_tumor Homo sapiens cDNA clone
            IMAGE:2748637 3', mRNA sequence.
ACCESSION   AW516596
VERSION     AW516596.1  GI:7154678
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 453)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 451.
FEATURES             Location/Qualifiers
     source          1. .453
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2748637"
                     /clone_lib="Soares_NHCeC_cervical_tumor"
                     /tissue_type="tumor"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: cervix; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCGAAGTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized; constructed by Bento Soares and
                     M.Fatima Bonaldo."
BASE COUNT       95 a     92 c    142 g    124 t
ORIGIN
```

```
  Query Match             55.8%;  Score 450;  DB 10;  Length 453;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-214;
  Matches  450;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    343 CTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGG 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    453 CTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGG 394

Qy    403 CTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAAC 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    393 CTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAAC 334

Qy    463 TTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCC 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 TTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCC 274

Qy    523 TCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGAC 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 TCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGAC 214

Qy    583 CATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAG 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 CATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAG 154

Qy    643 AGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGG 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 AGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGG 94

Qy    703 TCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACA 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 TCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACA 34

Qy    763 AACAAGTAATAAATAAATAAATAAACTCAA 792
          ||||||||||||||||||||||||||||||
Db     33 AACAAGTAATAAATAAATAAATAAACTCAA 4
```

```
RESULT 4
BE899580
LOCUS       BE899580                906 bp    mRNA    linear   EST 29-SEP-2000
DEFINITION  601682443F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3952523 5',
            mRNA sequence.
ACCESSION   BE899580
VERSION     BE899580.1  GI:10367234
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM822  row: f  column: 12
            High quality sequence start: 23
            High quality sequence stop: 775.
FEATURES             Location/Qualifiers
     source          1. .906
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3952523"
                     /clone_lib="NIH_MGC_9"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      219 a    282 c    219 g    186 t
ORIGIN
```

```
  Query Match             53.2%;  Score 429;  DB 12;  Length 906;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-203;
  Matches  429;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     57 TCTGTATCCTGCTTCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCA 116
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      99 TCTGTATCCTGCTTCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCA 158

Qy     117 AGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAA 176
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     159 AGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAA 218

Qy     177 CAAACCTGAAAGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCC 236
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     219 CAAACCTGAAAGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCC 278

Qy     237 GCCTTTGGGTGGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAG 296
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     279 GCCTTTGGGTGGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAG 338

Qy     297 ACAGCGGAGAACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTC 356
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     339 ACAGCGGAGAACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTC 398

Qy     357 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     399 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 458

Qy     417 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     459 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 518

Qy     477 TGAGCCAAC 485
           |||||||||
Db     519 TGAGCCAAC 527
```

```
RESULT 5
AW970357/c
LOCUS       AW970357                543 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST382438 MAGE resequences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW970357
VERSION     AW970357.1  GI:8160202
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 276
            Seq primer: Forward.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGK"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      120 a    121 c    151 g    151 t
ORIGIN

  Query Match           50.9%;  Score 410;  DB 10;  Length 543;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-194;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     380 ATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTC 439
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     410 ATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTC 351
```

```
Qy     440 CTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGG 499
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     350 CTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGG 291

Qy     500 CCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCG 559
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     290 CCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCG 231

Qy     560 GTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCA 619
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     230 GTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCA 171

Qy     620 AGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGC 679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     170 AGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGC 111

Qy     680 ATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGT 739
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110 ATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGT 51

Qy     740 CTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      50 CTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 1
```

```
RESULT 6
AI339648/c
LOCUS       AI339648                413 bp    mRNA    linear   EST 13-FEB-1999
DEFINITION  qk63a12.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1873630 3',
            mRNA sequence.
ACCESSION   AI339648
VERSION     AI339648.1  GI:4076575
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 846   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 407.
FEATURES             Location/Qualifiers
     source          1. .413
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1873630"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       82 a     82 c    128 g    121 t
```

```
ORIGIN

  Query Match            50.4%;  Score 406;  DB 9;  Length 413;
  Best Local Similarity  100.0%;  Pred. No. 3.7e-192;
  Matches  406;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     387 AGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCT 446
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     413 AGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCT 354

Qy     447 CAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGT 506
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     353 CAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGT 294

Qy     507 CGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAG 566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     293 CGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAG 234

Qy     567 CTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGG 626
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     233 CTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGG 174

Qy     627 CTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAAC 686
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173 CTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAAC 114

Qy     687 TACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGG 746
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     113 TACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGG 54

Qy     747 AAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTCAA 792
           |||||||||||||||||||||||||||||||||||||||||||||||
Db      53 AAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTCAA 8


RESULT 7
AI948903/c
LOCUS       AI948903                450 bp    mRNA    linear   EST 08-MAR-2000
DEFINITION  wq17c02.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2471522 3',
            mRNA sequence.
ACCESSION   AI948903
VERSION     AI948903.1  GI:5741213
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 450)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 607   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 438.
FEATURES             Location/Qualifiers
     source          1. .450
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2471522"
                     /clone_lib="NCI_CGAP_Kid12"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1323912-1325831, 1471368-1472903 and
                     1492104-1493255). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
BASE COUNT      94 a     92 c    141 g    123 t
ORIGIN

  Query Match            49.4%;  Score 398;  DB 9;  Length 450;
  Best Local Similarity  99.8%;  Pred. No. 3.8e-188;
  Matches  448;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     344 TGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGC 403
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     450 TGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGC 391

Qy     404 TGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACT 463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     390 TGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACT 331

Qy     464 TTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCT 523
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     330 TTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCT 271

Qy     524 CTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACC 583
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     270 CTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACC 211

Qy     584 ATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGA 643
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     210 ATCAATCCTGCTAGAGTGCAGGGAGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGA 151

Qy     644 GTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGT 703
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     150 GTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGT 91

Qy     704 CCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAA 763
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90 CCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAA 31

Qy     764 ACAAGTAATAAATAAATAAATAAACTCAA 792
           ||||||||||||||||||||||||||||
Db      30 ACAAGTAATAAATAAATAAATAAACTCAA 2


RESULT 8
AI833297/c
LOCUS       AI833297                444 bp    mRNA    linear   EST 13-JUL-1999
DEFINITION  at67a07.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
            IMAGE:2377044 3', mRNA sequence.
ACCESSION   AI833297
VERSION     AI833297.1  GI:5455277
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
```

```
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .444
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2377044"
                      /clone_lib="Barstead colon HPLRB7"
                      /sex="male"
                      /dev_stage="adult, age 25"
                      /lab_host="DH10B (phage resistant)"
                      /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      [5' AATTCACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested
                      with Not I and cloned into the Not I and Eco RI sites of
                      the modified pT7T3 vector. Library constructed by Bob
                      Barstead."
BASE COUNT       93 a     92 c    141 g    118 t
ORIGIN

  Query Match             48.8%;  Score 393;  DB 9;  Length 444;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-185;
  Matches  443;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    347 CTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGC 406
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    444 CTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGC 385

Qy    407 AAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTC 466
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    384 AAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTC 325

Qy    467 AGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTC 526
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    324 AGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTC 265

Qy    527 CAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATC 586
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    264 CAGTGCAAACCACCGAGCATTCCACCATGGCCGGTCACAGCTACAAATCCAGAGACCATC 205

Qy    587 AATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTC 646
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    204 AATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTC 145

Qy    647 TCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCC 706
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 TCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCC 85

Qy    707 CAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACA 766
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     84 CAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACA 25

Qy    767 AGTAATAAATAAATAAATAAACTC 790
          |||||||||||||||||||||||||
Db     24 AGTAATAAATAAATAAATAAACTC 1


RESULT 9
AI304380/c
LOCUS       AI304380                467 bp    mRNA    linear   EST 01-FEB-1999
DEFINITION  qo59c12.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1912822 3',
            mRNA sequence.
ACCESSION   AI304380
VERSION     AI304380.1  GI:3988069
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 467)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 604    Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 438.
FEATURES              Location/Qualifiers
     source           1. .467
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1912822"
                      /clone_lib="NCI_CGAP_Co8"
                      /tissue_type="adenocarcinoma"
                      /lab_host="DH10B"
                      /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; 1st strand cDNA was prepared from
                      colon adenocarcinoma, and was then primed with a Not I -
                      oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                      RI adaptors (Pharmacia), digested with Not I and cloned
                      into the Not I and Eco RI sites of the modified pT7T3
                      vector.  Library is normalized. Library was constructed by
                      Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       96 a     97 c    148 g    126 t
ORIGIN

  Query Match             47.1%;  Score 380;  DB 9;  Length 467;
  Best Local Similarity   99.8%;  Pred. No. 4e-179;
  Matches  430;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    325 GAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTT 384
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    467 GAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTT 408

Qy    385 GGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGT 444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    407 GGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGT 348

Qy    445 CTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCA 504
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    347 CTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCA 288

Qy    505 GTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCAC 564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    287 GTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCAC 228

Qy    565 AGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGT 624
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    227 AGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGAGGCAAGCACCCAAGGGT 168

Qy    625 GGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTA 684
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    167 GGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTA 108

Qy    685 ACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCT 744
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107 ACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCT 48

Qy    745 GGAAAGAGCAT 755
          ||||||||||||
```

```
Db     47 GGAAAGAGCAT 37


RESULT 10
AI283185/c
LOCUS       AI283185                430 bp    mRNA    linear   EST 28-JAN-1999
DEFINITION  qk49g09.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1872352 3',
            mRNA sequence.
ACCESSION   AI283185
VERSION     AI283185.1  GI:3921418
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 933   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 403.
FEATURES             Location/Qualifiers
     source          1. .430
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1872352"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       88 a     90 c    133 g    119 t
ORIGIN

  Query Match             47.0%;  Score 379;  DB 9;  Length 430;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-178;
  Matches  429;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     361 TTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 TTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAA 371

Qy     421 AGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 AGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAG 311

Qy     481 CCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     310 CCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACC 251

Qy     541 GAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 GAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGT 191

Qy     601 GCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 GCAGGGAGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAG 131

Qy     661 GTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTC 71

Qy     721 CTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAAT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 CTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAAT 11

Qy     781 AAATAAACTC 790
           ||||||||||
Db      10 AAATAAACTC 1


RESULT 11
AI813445/c
LOCUS       AI813445                429 bp    mRNA    linear   EST 21-DEC-1999
DEFINITION  wj06e01.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2402040 3',
            mRNA sequence.
ACCESSION   AI813445
VERSION     AI813445.1  GI:5424660
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 616   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 389.
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2402040"
                     /clone_lib="NCI_CGAP_Kid12"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1323912-1325831, 1471368-1472903 and
                     1492104-1493255). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
BASE COUNT       86 a     90 c    132 g    121 t
ORIGIN

  Query Match             46.8%;  Score 377;  DB 9;  Length 429;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-177;
  Matches  427;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     365 CAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTT 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    429 CAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTT 370

Qy    425 CCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAA 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    369 CCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAA 310

Qy    485 CCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGC 544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    309 CCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGC 250

Qy    545 ATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAG 604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 ATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAG 190

Qy    605 GGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCC 664
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 GGAGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCC 130

Qy    665 ATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAG 724
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 ATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAG 70

Qy    725 CCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAAT 784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 CCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAAT 10

Qy    785 AAACTCAA 792
          ||||||||
Db      9 AAACTCAA 2
```

```
RESULT 12
AW854263/c
LOCUS       AW854263                517 bp    mRNA    linear   EST 19-MAY-2000
DEFINITION  RC3-CT0254-100500-211-g12 CT0254 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW854263
VERSION     AW854263.1  GI:7949956
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 517)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-CT0254-100
            500-211-g12&t3=2000-05-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 6
            High quality sequence stop: 517.
FEATURES             Location/Qualifiers
     source          1. .517
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="CT0254"
                     /dev_stage="Adult"
                     /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT      106 a    118 c    172 g    121 t
ORIGIN
```

```
  Query Match             46.0%;  Score 371;  DB 10;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-174;
  Matches  371;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    331 CCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGG 390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 CCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGG 375

Qy    391 CTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAAT 450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 CTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAAT 315

Qy    451 TGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCC 510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 TGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCC 255

Qy    511 ACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTAC 570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 ACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTAC 195

Qy    571 AAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGA 630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 AAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGA 135

Qy    631 CCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACC 690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134 CCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACC 75

Qy    691 AGCATCCAGTG 701
          |||||||||||
Db     74 AGCATCCAGTG 64
```

```
RESULT 13
AA938765/c
LOCUS       AA938765                420 bp    mRNA    linear   EST 23-JUN-1998
DEFINITION  on44h11.s1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1559589 3',
            mRNA sequence.
ACCESSION   AA938765
VERSION     AA938765.1  GI:3096793
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 420)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 613   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
```

```
            High quality sequence stop: 403.
FEATURES             Location/Qualifiers
     source          1. .420
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1559589"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       86 a     88 c    130 g    116 t
ORIGIN

  Query Match            45.8%;  Score 369;  DB 9;  Length 420;
  Best Local Similarity  99.8%;  Pred. No. 1.2e-173;
  Matches  419;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    371 TGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAA 430
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 TGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAA 361

Qy    431 CTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCAC 490
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 CTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCAC 301

Qy    491 CCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCA 550
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 CCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCA 241

Qy    551 CCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGC 610
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    240 CCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGAGGC 181

Qy    611 AAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAG 670
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 AAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAG 121

Qy    671 CCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCT 730
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCT 61

Qy    731 GACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTC 790
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTC 1


RESULT 14
AI832391/c
LOCUS       AI832391                455 bp    mRNA    linear   EST 13-JUL-1999
DEFINITION  at68f03.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
            IMAGE:2377181 3', mRNA sequence.
ACCESSION   AI832391
VERSION     AI832391.1  GI:5454380
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 455)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 420.
FEATURES             Location/Qualifiers
     source          1. .455
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2377181"
                     /clone_lib="Barstead colon HPLRB7"
                     /sex="male"
                     /dev_stage="adult, age 25"
                     /lab_host="DH10B (phage resistant)"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [5' AATTCACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested
                     with Not I and cloned into the Not I and Eco RI sites of
                     the modified pT7T3 vector. Library constructed by Bob
                     Barstead."
BASE COUNT       97 a     94 c    145 g    119 t
ORIGIN

  Query Match            45.4%;  Score 366;  DB 9;  Length 455;
  Best Local Similarity  99.8%;  Pred. No. 4e-172;
  Matches  416;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    341 TCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCG 400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    455 TCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCG 396

Qy    401 GGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCA 460
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    395 GGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCGCGTCCTTGTCTCAATTGTGCCATCA 336

Qy    461 ACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGG 520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    335 ACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGG 276

Qy    521 CCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAG 580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    275 CCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAG 216

Qy    581 ACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGC 640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    215 ACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGC 156

Qy    641 AGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGT 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    155 AGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGT 96

Qy    701 GGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCC 757
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     95 GGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCC 39


RESULT 15
AI831407/c
LOCUS       AI831407                467 bp    mRNA    linear   EST 21-DEC-1999
DEFINITION  wj64a12.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2407582 3',
            mRNA sequence.
ACCESSION   AI831407
VERSION     AI831407.1  GI:5452078
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 467)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 756   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 458.
FEATURES              Location/Qualifiers
     source           1. .467
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2407582"
                      /clone_lib="NCI_CGAP_Lu19"
                      /tissue_type="squamous cell carcinoma, poorly
                      differentiated (4 pooled tumors, including primary and
                      metastatic)"
                      /dev_stage="adult"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; 1st strand cDNA was prepared from
                      pooled lung tumor tissue, and was then primed with a Not I
                      - oligo(dT) primer. Double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I  and Eco RI sites of the modified
                      pT7T3 vector.  Library went througb one round of
                      normalization. Library constructed by Bento Soares and M.
                      Fatima Bonaldo. "
BASE COUNT       95 a     94 c    146 g    132 t
ORIGIN

  Query Match             44.9%;  Score 362;  DB 9;  Length 467;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-170;
  Matches  462;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    343 CTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGG 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    467 CTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGG 408

Qy    403 CTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAAC 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    407 CTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAAC 348

Qy    463 TTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCC 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    347 TTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCC 288

Qy    523 TCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGAC 582
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    287 TCTCCAGTGCAAACCCCCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGAC 228

Qy    583 CATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAG 642
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    227 CATCAATCCTGCTAGAGTGCAGGGAGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAG 168

Qy    643 AGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGG 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    167 AGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGG 108

Qy    703 TCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACA 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    107 TCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACA 48

Qy    763 AACAAGTAATAAATAAATAAATAAACTCAATGCAGACACAAAAA 806
          |||||||||||||||||||||||||||||||||||||||||||||
Db     47 AACAAGTAATAAATAAATAAATAAACTCAATGCAGACACAAAAA 4


Search completed: May 17, 2003, 17:12:07
Job time : 1749 secs
```

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 13:37:01 ; Search time 56 Seconds
(without alignments)
4413.955 Million cell updates/sec

Title: US-09-724-000A-4
Perfect score: 806
Sequence: 1 ggaacgagggaaaatctgcc.....actcaatgcagacacaaaaa 806

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	2.5	12847	1	US-08-550-715-1	Sequence 1, Appli
c	2	19	2.4	425	4	US-09-641-638-326	Sequence 326, App
	3	19	2.4	465	6	5496550-9	Patent No. 5496550
	4	19	2.4	466	6	5496550-7	Patent No. 5496550
c	5	19	2.4	472	4	US-08-943-731-123	Sequence 123, App
	6	19	2.4	2249	2	US-08-463-081B-34	Sequence 34, Appl
	7	19	2.4	2249	2	US-08-461-379A-34	Sequence 34, Appl
	8	19	2.4	2249	2	US-08-462-390B-34	Sequence 34, Appl
	9	19	2.4	2249	3	US-08-463-074B-34	Sequence 34, Appl
	10	19	2.4	2249	3	US-08-465-585C-34	Sequence 34, Appl
	11	19	2.4	2249	3	US-08-652-446-34	Sequence 34, Appl
c	12	19	2.4	2318	4	US-09-851-062-3	Sequence 3, Appli
c	13	19	2.4	2362	1	US-08-265-087-1	Sequence 1, Appli
c	14	19	2.4	2362	1	US-08-621-493-1	Sequence 1, Appli
c	15	19	2.4	2362	2	US-08-965-688-1	Sequence 1, Appli
c	16	19	2.4	2362	4	US-09-260-173-1	Sequence 1, Appli
	17	19	2.4	3865	1	US-08-832-883-48	Sequence 48, Appl
	18	19	2.4	3865	2	US-08-832-877-48	Sequence 48, Appl
	19	19	2.4	4281	4	US-09-357-206A-8	Sequence 8, Appli
	20	19	2.4	5630	2	US-08-937-931-1	Sequence 1, Appli
	21	19	2.4	5630	4	US-09-285-502-1	Sequence 1, Appli
	22	19	2.4	5630	4	US-09-709-126-1	Sequence 1, Appli
	23	19	2.4	5630	4	US-09-871-385A-1	Sequence 1, Appli
c	24	19	2.4	7521	4	US-09-004-838-116	Sequence 116, App
c	25	19	2.4	7874	4	US-09-780-175-96	Sequence 96, Appl
c	26	19	2.4	9365	4	US-09-608-285A-8	Sequence 8, Appli
c	27	19	2.4	9365	4	US-09-350-836B-8	Sequence 8, Appli
c	28	19	2.4	9365	4	US-09-370-265-8	Sequence 8, Appli
	29	19	2.4	10881	4	US-09-357-206A-9	Sequence 9, Appli
	30	19	2.4	12286	4	US-09-357-206A-1	Sequence 1, Appli
c	31	19	2.4	14747	4	US-09-608-285A-42	Sequence 42, Appl
	32	19	2.4	14796	4	US-08-975-080-35	Sequence 35, Appl
	33	19	2.4	14796	4	US-09-630-706-10	Sequence 10, Appl
	34	19	2.4	14796	4	US-09-496-694B-3	Sequence 3, Appli
c	35	19	2.4	15977	4	US-09-608-285A-59	Sequence 59, Appl
	36	19	2.4	18073	4	US-09-078-294-12	Sequence 12, Appl
c	37	19	2.4	24183	4	US-08-943-731-3	Sequence 3, Appli
	38	19	2.4	72604	4	US-09-268-992-7	Sequence 7, Appli
	39	19	2.4	72604	4	US-09-657-474-7	Sequence 7, Appli
	40	19	2.4	98844	4	US-09-791-211-10	Sequence 10, Appl
c	41	19	2.4	162450	4	US-09-345-882-1	Sequence 1, Appli
c	42	18	2.2	21	4	US-09-165-868-8	Sequence 8, Appli
c	43	18	2.2	46	1	US-08-222-177A-349	Sequence 349, App
	44	18	2.2	66	1	US-08-222-177A-301	Sequence 301, App
	45	18	2.2	206	2	US-08-623-906A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-550-715-1
; Sequence 1, Application US/08550715
; Patent No. 5750345
; GENERAL INFORMATION:
; APPLICANT: Bowie, Lemuel J.
; TITLE OF INVENTION: Human '-Thalassemia Mutations as a Predictor of
; TITLE OF INVENTION: Blood-Related Disorders
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/550,715
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28493/32834
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12847 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(6703..6797, 6915..7119, 7262..7387)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(10514..10608, 10726..10930, 11080..11205)
US-08-550-715-1

Query Match 2.5%; Score 20; DB 1; Length 12847;
Best Local Similarity 100.0%; Pred. No. 1.8;

```
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    770 AATAAATAAATAAATAAACT 789
          ||||||||||||||||||||
Db   4590 AATAAATAAATAAATAAACT 4609


RESULT 2
US-09-641-638-326/c
; Sequence 326, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 326
;   LENGTH: 425
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 276
;   OTHER INFORMATION: 10-65-276 : polymorphic base A or G
;   NAME/KEY: misc_binding
;   LOCATION: 256..275
;   OTHER INFORMATION: 10-65-276.mis1, potential
;   NAME/KEY: misc_binding
;   LOCATION: 277..295
;   OTHER INFORMATION: 10-65-276.mis2, complement
;   NAME/KEY: primer_bind
;   LOCATION: 1..19
;   OTHER INFORMATION: upstream amplification primer
;   NAME/KEY: primer_bind
;   LOCATION: 408..425
;   OTHER INFORMATION: downstream amplification primer, complement
;   NAME/KEY: misc_binding
;   LOCATION: 264..288
;   OTHER INFORMATION: 10-65-276 potential probe
US-09-641-638-326

  Query Match             2.4%;  Score 19;  DB 4;  Length 425;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    770 AATAAATAAATAAATAAAC 788
          ||||||||||||||||||
Db    350 AATAAATAAATAAATAAAC 332


RESULT 3
5496550-9
;Patent No. 5496550
;    APPLICANT: WALLACH, MICHAEL;PUGATSCH, THEA;MENCHER, DAVID
;    TITLE OF INVENTION: METHOD OF REDUCING THE OUTPUT OF EIMERIA
;OOCYSTS FROM A NEWBORN CHICK
;    NUMBER OF SEQUENCES: 10
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/108,763
;      FILING DATE: 17-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 642,219
;      FILING DATE: 16-JAN-1991
;      APPLICATION NUMBER: 310,603
;      FILING DATE: 14-FEB-1989
;      APPLICATION NUMBER: 155,245
;      FILING DATE: 12-FEB-1988
;      APPLICATION NUMBER: 896,611
;      FILING DATE: 14-AUG-1986
;SEQ ID NO:9:
;      LENGTH: 465
5496550-9

  Query Match             2.4%;  Score 19;  DB 6;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    769 TAATAAATAAATAAATAAA 787
          |||||||||||||||||||
Db    203 TAATAAATAAATAAATAAA 221


RESULT 4
5496550-7
;Patent No. 5496550
;    APPLICANT: WALLACH, MICHAEL;PUGATSCH, THEA;MENCHER, DAVID
;    TITLE OF INVENTION: METHOD OF REDUCING THE OUTPUT OF EIMERIA
;OOCYSTS FROM A NEWBORN CHICK
;    NUMBER OF SEQUENCES: 10
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/108,763
;      FILING DATE: 17-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 642,219
;      FILING DATE: 16-JAN-1991
;      APPLICATION NUMBER: 310,603
;      FILING DATE: 14-FEB-1989
;      APPLICATION NUMBER: 155,245
;      FILING DATE: 12-FEB-1988
;      APPLICATION NUMBER: 896,611
;      FILING DATE: 14-AUG-1986
;SEQ ID NO:7:
;      LENGTH: 466
5496550-7

  Query Match             2.4%;  Score 19;  DB 6;  Length 466;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    769 TAATAAATAAATAAATAAA 787
          |||||||||||||||||||
Db    204 TAATAAATAAATAAATAAA 222


RESULT 5
US-08-943-731-123/c
; Sequence 123, Application US/08943731
; Patent No. 6265157
;  GENERAL INFORMATION:
;    APPLICANT:  PROCKOP, DARWIN J.
;    APPLICANT:  SPOTILA, LORETTA D.
;    APPLICANT:  DELTAS, CONSTANTINOS D.
;    APPLICANT:  SEREDA, LARISA
;    APPLICANT:  LARSON, ANDREA W.
;    APPLICANT:  PACK, MICHAEL
;    APPLICANT:  COLIGE, ALAIN
;    APPLICANT:  EARLY, JAMES
;    APPLICANT:  KORKKO, JARMO
;    APPLICANT:  ALA-KOKKO, LEENA, et al.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR DETECTING
;    TITLE OF INVENTION:  ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
```

```
;    NUMBER OF SEQUENCES:  666
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PANITCH SCHWARZE JACOBS & NADEL, P.C.
;      STREET:  ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
;      STREET:  FLR.
;      CITY:  PHILADELPHIA
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-7086
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,731
;      FILING DATE:  03-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,322
;      FILING DATE:  14-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/803,628
;      FILING DATE:  03-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOYLE LEARY Ph.D., KATHRYN
;      REGISTRATION NUMBER:  36,317
;      REFERENCE/DOCKET NUMBER:  9598-27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-965-1284
;      TELEFAX:  215-567-2991
;      TELEX:  831-494
;  INFORMATION FOR SEQ ID NO:  123:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  472 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-943-731-123

  Query Match             2.4%;  Score 19;  DB 4;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     769 TAATAAATAAATAAATAAA 787
           |||||||||||||||||||
Db     106 TAATAAATAAATAAATAAA 88


RESULT 6
US-08-463-081B-34
; Sequence 34, Application US/08463081B
; Patent No. 5871960
; Patent No. 5871960 5837487
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR5 Polypeptide,
;    TITLE OF INVENTION:  Vector and Transformed Cell Thereof, and Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      STREET:  444 South Flower St. - Suite 1900
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,081B
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/104,736
;      FILING DATE:  10-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38150 (DART-060)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 489-4210
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2249 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-463-081B-34

  Query Match             2.4%;  Score 19;  DB 2;  Length 2249;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     770 AATAAATAAATAAATAAAC 788
           |||||||||||||||||||
Db    1817 AATAAATAAATAAATAAAC 1835


RESULT 7
US-08-461-379A-34
; Sequence 34, Application US/08461379A
; Patent No. 5871961
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR5 Polypeptide,
;    TITLE OF INVENTION:  Vector and Transformed Cell Thereof, and
;    TITLE OF INVENTION:  Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      ADDRESSEE:  .          (B) STREET:One Westlakes-Berwyn
;      CITY:  Valley Forge
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,379A
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108; 08/104,736
;      APPLICATION NUMBER:  & 07/796,066
;      FILING DATE:  27-OCT-1994; 10-AUG-1993 & 20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  DART-070
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)470-0700
;      TELEFAX:  (610)470-0701
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  2249 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-461-379A-34

  Query Match             2.4%;  Score 19;  DB 2;  Length 2249;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    770 AATAAATAAATAAATAAAC 788
          ||||||||||||||||||||
Db   1817 AATAAATAAATAAATAAAC 1835


RESULT 8
US-08-462-390B-34
; Sequence 34, Application US/08462390B
; Patent No. 5882894
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, K. A., & Beadling, C.
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR8 Polypeptide, Vector and
;    TITLE OF INVENTION:  Transformed Cell Thereof, and Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      ADDRESSEE:  .          (B) STREET:One Westlakes-Berwyn
;      CITY:  Valley Forge
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,390B
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108
;      FILING DATE:  27-OCT-1994
;      APPLICATION NUMBER:  USSN 08/104,736
;      FILING DATE:  10-AUG-1993
;      APPLICATION NUMBER:  USSN 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  DART-040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)407-0700
;      TELEFAX:  (610)407-0701
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2249 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-462-390B-34

  Query Match             2.4%;  Score 19;  DB 2;  Length 2249;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    770 AATAAATAAATAAATAAAC 788
          ||||||||||||||||||||
Db   1817 AATAAATAAATAAATAAAC 1835


RESULT 9
US-08-463-074B-34
; Sequence 34, Application US/08463074B
; Patent No. 6020155
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR1 Fusion Protein, Vector an
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      ADDRESSEE:  .          (B) STREET:              444 South Flower St. - Suite 190
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,074B
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/104,736
;      FILING DATE:  10-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38143 (DART-020)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 489-4210
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2249 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-463-074B-34

  Query Match             2.4%;  Score 19;  DB 3;  Length 2249;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    770 AATAAATAAATAAATAAAC 788
          ||||||||||||||||||||
Db   1817 AATAAATAAATAAATAAAC 1835


RESULT 10
US-08-465-585C-34
; Sequence 34, Application US/08465585C
; Patent No. 6027914
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, K. A., & Beadling, C.
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR6 Polypeptide, Vecto
;    TITLE OF INVENTION:  Transformed Cell Thereof, and Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY,SCHROEDER & POPLAWSKI
;      ADDRESSEE:  .          (B) STREET:              444South Flower St. - Suite 19
;      CITY:  Los Angeles
;      STATE:  Californiaa
;      COUNTRY:  USA
;      ZIP:  900071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,585C
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108
;      FILING DATE:  27-OCT-1994
;      APPLICATION NUMBER:  USSN 08/104,736
;      FILING DATE:  10-AUG-1993
;      APPLICATION NUMBER:  USSN 07/796,066
;      FILING DATE:  20-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38149 (DART-050)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 4894210
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2249 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-465-585C-34

  Query Match             2.4%;  Score 19;  DB 3;  Length 2249;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    770 AATAAATAAATAAATAAAC 788
          ||||||||||||||||||||
Db   1817 AATAAATAAATAAATAAAC 1835


RESULT 11
US-08-652-446-34
; Sequence 34, Application US/08652446
; Patent No. 6057427
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR5
;    TITLE OF INVENTION:  Polypeptide, Vector and Transformed Cell Thereof, and
;    TITLE OF INVENTION:  Expression Thereof
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      ADDRESSEE:  .          (B) STREET:              444 South Flower St. - Suite 1900
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,446
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP App. # 96921319.8
;      FILING DATE:  5-JAN-1998
;      APPLICATION NUMBER:  PCT/US/96/09194
;      FILING DATE:  5-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/330,108
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/463,074
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/462;337
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/462,390
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/465,585
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/463,081
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/461,379
;      FILING DATE:  5-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/739,523
;      FILING DATE:  29-OCT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  FP66 40035
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 489-4210
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2249 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-652-446-34

  Query Match             2.4%;  Score 19;  DB 3;  Length 2249;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    770 AATAAATAAATAAATAAAC 788
          ||||||||||||||||||||
Db   1817 AATAAATAAATAAATAAAC 1835


RESULT 12
US-09-851-062-3/c
; Sequence 3, Application US/09851062
; Patent No. 6448081
; GENERAL INFORMATION:
;  APPLICANT: Brenda F. Baker
;  APPLICANT:  Susan M. Freier
;  TITLE OF INVENTION: ANTISENSE MODULATION OF INTERLEUKIN 12 P40 SUBUNIT EXPRESSION
;  FILE REFERENCE: RTS-0247
;  CURRENT APPLICATION NUMBER: US/09/851,062
;  CURRENT FILING DATE:  2001-05-07
;  NUMBER OF SEQ ID NOS: 87
; SEQ ID NO 3
;   LENGTH: 2318
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (14)...(1000)
US-09-851-062-3

  Query Match             2.4%;  Score 19;  DB 4;  Length 2318;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    770 AATAAATAAATAAATAAAC 788
          ||||||||||||||||||||
Db   2274 AATAAATAAATAAATAAAC 2256
```

```
RESULT 13
US-08-265-087-1/c
; Sequence 1, Application US/08265087
; Patent No. 5571515
;  GENERAL INFORMATION:
;    APPLICANT:  Scott, Phillip
;    APPLICANT:  Trinchieri, Giorgio
;    TITLE OF INVENTION:  Compositions and Methods for Use of
;    TITLE OF INVENTION:  IL-12 as an Adjuvant
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Center, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/265,087
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/229,282
;      FILING DATE:  18-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST51AUSA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9206
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2362 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  33..1016
US-08-265-087-1

  Query Match             2.4%;  Score 19;  DB 1;  Length 2362;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     770 AATAAATAAATAAATAAAC 788
           ||||||||||||||||||||
Db    2293 AATAAATAAATAAATAAAC 2275


RESULT 14
US-08-621-493-1/c
; Sequence 1, Application US/08621493
; Patent No. 5723127
;  GENERAL INFORMATION:
;    APPLICANT:  Scott, Phillip
;    APPLICANT:  Trinchieri, Giorgio
;    TITLE OF INVENTION:  Compositions and Methods for Use of
;    TITLE OF INVENTION:  IL-12 as an Adjuvant
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Center, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,493
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/265,087
;      FILING DATE:  17-JUN-1994
;      APPLICATION NUMBER:  US 08/229,282
;      FILING DATE:  18-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST51AUSA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9206
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2362 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  33..1016
US-08-621-493-1

  Query Match             2.4%;  Score 19;  DB 1;  Length 2362;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     770 AATAAATAAATAAATAAAC 788
           ||||||||||||||||||||
Db    2293 AATAAATAAATAAATAAAC 2275


RESULT 15
US-08-965-688-1/c
; Sequence 1, Application US/08965688
; Patent No. 5976539
;  GENERAL INFORMATION:
;    APPLICANT:  Scott, Phillip
;    APPLICANT:  Trinchieri, Giorgio
;    TITLE OF INVENTION:  Compositions and Methods for Use of
;    TITLE OF INVENTION:  IL-12 as an Adjuvant
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Center, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/965,688
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/621,493
;      FILING DATE:  25-MAR-1996
```

```
;       APPLICATION NUMBER:  08/265,087
;       FILING DATE:  17-JUN-1994
;       APPLICATION NUMBER:  US 08/229,282
;       FILING DATE:  18-APR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bak, Mary E.
;       REGISTRATION NUMBER:  31,215
;       REFERENCE/DOCKET NUMBER:  WST51AUSA
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  215-540-9206
;       TELEFAX:  215-540-5818
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2362 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  33..1016
US-08-965-688-1

  Query Match             2.4%;  Score 19;  DB 2;  Length 2362;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     770 AATAAATAAATAAATAAAC 788
           ||||||||||||||||||
Db    2293 AATAAATAAATAAATAAAC 2275


Search completed: May 17, 2003, 17:13:39
Job time : 86 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 14:37:12 ; Search time 116 Seconds
(without alignments)
8962.846 Million cell updates/sec

Title: US-09-724-000A-4
Perfect score: 806
Sequence: 1 ggaacgagggaaaatctgcc.....actcaatgcagacacaaaaa 806

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 810007 seqs, 644969091 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	783	97.1	800	9	US-09-981-353-177	Sequence 177, App
2	782	97.0	804	9	US-10-028-072-149	Sequence 149, App
3	782	97.0	804	9	US-10-121-049-149	Sequence 149, App
4	782	97.0	804	9	US-10-123-904-149	Sequence 149, App
5	782	97.0	804	9	US-10-140-470-149	Sequence 149, App
6	782	97.0	804	9	US-10-175-746-149	Sequence 149, App
7	782	97.0	804	9	US-10-176-918-149	Sequence 149, App
8	782	97.0	804	9	US-10-176-921-149	Sequence 149, App
9	782	97.0	804	9	US-10-137-865-149	Sequence 149, App
10	782	97.0	804	9	US-10-140-474-149	Sequence 149, App
11	782	97.0	804	9	US-10-142-431-149	Sequence 149, App
12	782	97.0	804	9	US-10-143-114-149	Sequence 149, App
13	782	97.0	804	9	US-10-140-002-149	Sequence 149, App
14	782	97.0	804	9	US-10-142-419-149	Sequence 149, App
15	782	97.0	804	9	US-10-123-262-149	Sequence 149, App
16	782	97.0	804	9	US-10-142-423-149	Sequence 149, App
17	782	97.0	804	9	US-10-121-050-149	Sequence 149, App
18	782	97.0	804	9	US-10-141-755-149	Sequence 149, App
19	782	97.0	804	9	US-10-143-032-149	Sequence 149, App
20	782	97.0	804	9	US-10-123-108-149	Sequence 149, App
21	782	97.0	804	9	US-10-123-236-149	Sequence 149, App
22	782	97.0	804	9	US-10-123-261-149	Sequence 149, App
23	782	97.0	804	9	US-10-140-921-149	Sequence 149, App
24	782	97.0	804	9	US-10-140-928-149	Sequence 149, App
25	782	97.0	804	9	US-10-121-045-149	Sequence 149, App
26	782	97.0	804	9	US-10-123-292-149	Sequence 149, App
27	782	97.0	804	9	US-10-123-903-149	Sequence 149, App
28	782	97.0	804	9	US-10-124-819-149	Sequence 149, App
29	782	97.0	804	9	US-10-124-822-149	Sequence 149, App
30	782	97.0	804	9	US-10-140-925-149	Sequence 149, App
31	782	97.0	804	9	US-10-160-498-149	Sequence 149, App
32	782	97.0	804	9	US-10-121-041-149	Sequence 149, App
33	782	97.0	804	9	US-10-121-043-149	Sequence 149, App
34	782	97.0	804	9	US-10-121-047-149	Sequence 149, App
35	782	97.0	804	9	US-10-123-215-149	Sequence 149, App
36	782	97.0	804	9	US-10-123-902-149	Sequence 149, App
37	782	97.0	804	9	US-10-123-908-149	Sequence 149, App
38	782	97.0	804	9	US-10-123-909-149	Sequence 149, App
39	782	97.0	804	9	US-10-123-910-149	Sequence 149, App
40	782	97.0	804	9	US-10-124-813-149	Sequence 149, App
41	782	97.0	804	9	US-10-124-817-149	Sequence 149, App
42	782	97.0	804	9	US-10-124-824-149	Sequence 149, App
43	782	97.0	804	9	US-10-125-922-149	Sequence 149, App
44	782	97.0	804	9	US-10-125-924-149	Sequence 149, App
45	782	97.0	804	9	US-10-127-825A-149	Sequence 149, App

ALIGNMENTS

RESULT 1
US-09-981-353-177
; Sequence 177, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 177
; LENGTH: 800
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1736965CB1
US-09-981-353-177

Query Match 97.1%; Score 783; DB 9; Length 800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 783; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 65

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 125

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 185

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 245
```

```
Qy    250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 305

Qy    310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 365

Qy    370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 425

Qy    430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 485

Qy    490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 545

Qy    550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 605

Qy    610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    606 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 665

Qy    670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    666 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 725

Qy    730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 785

Qy    790 CAA 792
          |||
Db    786 CAA 788
```

RESULT 2
US-10-028-072-149
; Sequence 149, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694

; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081695
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081818
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082999
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 19/98-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match 97.0%; Score 782; DB 9; Length 804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682
```

```
Qy    670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy    730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy    790 CA 791
          ||
Db    803 CA 804


RESULT 3
US-10-121-049-149
; Sequence 149, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C17
;  CURRENT APPLICATION NUMBER: US/10/121,049
;  CURRENT FILING DATE:  2002-04-12
;       Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-121-049-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy     70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy    130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy    190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy    250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy    310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy    370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy    430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy    490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy    550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy    610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy    670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy    730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy    790 CA 791
          ||
Db    803 CA 804


RESULT 4
US-10-123-904-149
; Sequence 149, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C54
;  CURRENT APPLICATION NUMBER: US/10/123,904
;  CURRENT FILING DATE:  2002-04-16
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-123-904-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
```

```
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy     730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy     790 CA 791
           ||
Db     803 CA 804


RESULT 5
US-10-140-470-149
; Sequence 149, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C160
;  CURRENT APPLICATION NUMBER: US/10/140,470
;  CURRENT FILING DATE:  2002-05-06
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-140-470-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742
```

```
Qy    730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy    790 CA 791
          ||
Db    803 CA 804


RESULT 6
US-10-175-746-149
; Sequence 149, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C353
;  CURRENT APPLICATION NUMBER: US/10/175,746
;  CURRENT FILING DATE:  2002-06-19
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-175-746-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy     70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy    130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy    190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy    250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy    310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy    370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy    430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy    490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy    550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy    610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy    670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy    730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy    790 CA 791
          ||
Db    803 CA 804


RESULT 7
US-10-176-918-149
; Sequence 149, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C382
;  CURRENT APPLICATION NUMBER: US/10/176,918
;  CURRENT FILING DATE:  2002-06-20
;     Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-176-918-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy     730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy     790 CA 791
           ||
Db     803 CA 804


RESULT 8
US-10-176-921-149
; Sequence 149, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C288
;  CURRENT APPLICATION NUMBER: US/10/176,921
;  CURRENT FILING DATE:  2002-06-20
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-176-921-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy     730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802
```

```
Qy    790 CA 791
          ||
Db    803 CA 804


RESULT 9
US-10-137-865-149
; Sequence 149, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C154
;  CURRENT APPLICATION NUMBER: US/10/137,865
;  CURRENT FILING DATE:  2002-05-03
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-137-865-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy     70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy    130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy    190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy    250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy    310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy    370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy    430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy    490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy    550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy    610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy    670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy    730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy    790 CA 791
          ||
Db    803 CA 804


RESULT 10
US-10-140-474-149
; Sequence 149, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C162
;  CURRENT APPLICATION NUMBER: US/10/140,474
;  CURRENT FILING DATE:  2002-05-06
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-140-474-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy     70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy     730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy     790 CA 791
           ||
Db     803 CA 804


RESULT 11
US-10-142-431-149
; Sequence 149, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C251
;  CURRENT APPLICATION NUMBER: US/10/142,431
;  CURRENT FILING DATE:  2002-05-10
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-142-431-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy     730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy     790 CA 791
           ||
Db     803 CA 804
```

```
RESULT 12
US-10-143-114-149
; Sequence 149, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C211
;  CURRENT APPLICATION NUMBER: US/10/143,114
;  CURRENT FILING DATE:  2002-05-09
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-143-114-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy     730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy     790 CA 791
           ||
Db     803 CA 804


RESULT 13
US-10-140-002-149
; Sequence 149, Application US/10140002
; Publication No. US20030037623A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C59
;  CURRENT APPLICATION NUMBER: US/10/140,002
;  CURRENT FILING DATE:  2002-05-06
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-140-002-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy    190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy    250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy    310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy    370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy    430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy    490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy    550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy    610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy    670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy    730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy    790 CA 791
          ||
Db    803 CA 804
```

```
RESULT 14
US-10-142-419-149
; Sequence 149, Application US/10142419
; Publication No. US20030044945A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C244
;  CURRENT APPLICATION NUMBER: US/10/142,419
;  CURRENT FILING DATE:  2002-05-10
;  Prior Apploication removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-142-419-149

  Query Match             97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy     70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy    130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy    190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy    250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy    310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy    370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy    430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy    490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy    550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy    610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy    670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy    730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy    790 CA 791
          ||
Db    803 CA 804
```

```
RESULT 15
US-10-123-262-149
; Sequence 149, Application US/10123262
```

```
; Publication No. US20030049816A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C38
;  CURRENT APPLICATION NUMBER: US/10/123,262
;  CURRENT FILING DATE:  2002-04-15
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-123-262-149

  Query Match            97.0%;  Score 782;  DB 9;  Length 804;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy     730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy     790 CA 791
           ||
Db     803 CA 804


Search completed: May 17, 2003, 17:20:47
Job time : 121 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 11:55:46 ; Search time 2635 Seconds
(without alignments)
8902.035 Million cell updates/sec

Title: US-09-724-000A-4
Perfect score: 806
Sequence: 1 ggaacgagggaaaatctgcc.....actcaatgcagacacaaaaa 806

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		806	100.0	806	6	AX342218	AX342218 Sequence
2		782	97.0	804	6	AX464016	AX464016 Sequence
3		616	76.4	198481	9	AC022389	AC022389 Homo sapi
4	c	537	66.6	801	6	AX027767	AX027767 Sequence
5	c	498	61.8	797	6	AX027773	AX027773 Sequence
6		480	59.6	2063	9	AK025416	AK025416 Homo sapi
7		323	40.1	485	6	AX351259	AX351259 Sequence
8	c	299	37.1	382	6	AX340424	AX340424 Sequence
9	c	196	24.3	366	6	AX330090	AX330090 Sequence
10	c	196	24.3	366	6	AX335580	AX335580 Sequence
11	c	196	24.3	366	6	AX408291	AX408291 Sequence
12		93	11.5	4159	6	AX342222	AX342222 Sequence
13	c	27	3.3	180176	10	AL671190	AL671190 Mouse DNA
14	c	25	3.1	93165	9	AL357974	AL357974 Human DNA
15	c	25	3.1	152245	2	AC022829	AC022829 Homo sapi
16	c	25	3.1	153311	2	AC107297	AC107297 Homo sapi
17	c	25	3.1	172043	9	AC078828	AC078828 Homo sapi
18	c	25	3.1	193903	2	AC091156	AC091156 Homo sapi
19	c	25	3.1	198057	2	AC073719	AC073719 Mus muscu
20		25	3.1	231443	2	AC073691	AC073691 Mus muscu
21		25	3.1	246962	2	AC073811	AC073811 Mus muscu
22	c	24	3.0	94737	2	AC126011	AC126011 Medicago
23	c	24	3.0	144818	9	AL139343	AL139343 Human DNA
24		24	3.0	157758	9	AC098824	AC098824 Homo sapi
25		24	3.0	162486	2	AC044779	AC044779 Homo sapi
26		24	3.0	165114	2	AC098323	AC098323 Rattus no
27		24	3.0	174194	2	AL845359	AL845359 Danio rer
28		24	3.0	180718	2	AL844538	AL844538 Mus muscu
29	c	24	3.0	192519	10	MMU278435	AJ278435 Mus muscu
30		24	3.0	202095	2	AC130044	AC130044 Rattus no
31		24	3.0	215516	2	AC024694	AC024694 Mus muscu
32	c	24	3.0	221647	10	AL591003	AL591003 Mouse DNA
33		24	3.0	224446	2	AC122109	AC122109 Mus muscu
34	c	23	2.9	7355	6	AX344889	AX344889 Sequence
35		23	2.9	57276	2	AC124271	AC124271 Homo sapi
36	c	23	2.9	57276	2	AC124271	AC124271 Homo sapi
37		23	2.9	59438	2	AC090314	AC090314 Homo sapi
38		23	2.9	64886	2	AC101145	AC101145 Mus muscu
39	c	23	2.9	64886	2	AC101145	AC101145 Mus muscu
40	c	23	2.9	66969	9	AL161789	AL161789 Human DNA
41	c	23	2.9	72259	2	AC100677	AC100677 Mus muscu
42	c	23	2.9	76396	2	AC012995	AC012995 Drosophil
43		23	2.9	80922	2	AC097830	AC097830 Rattus no
44		23	2.9	86945	2	AC002490	AC002490 Homo sapi
45	c	23	2.9	88806	9	AL603890	AL603890 Human DNA

ALIGNMENTS

RESULT 1
AX342218
LOCUS AX342218 806 bp DNA linear PAT 12-JAN-2002
DEFINITION Sequence 4 from Patent WO0198497.
ACCESSION AX342218
VERSION AX342218.1 GI:18151763
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Polverino,A.J. and Luethy,R.
TITLE Secreted epithelial colon stromal-1 polypeptides, nucleic acids
encoding the same and uses thereof

```
  JOURNAL   Patent: WO 0198497-A 4 27-DEC-2001;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .806
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             29. .274
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD20706.1"
                     /db_xref="GI:18151764"
                     /translation="MRLLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRV
                     PSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPGALPQV"
     sig_peptide     29. .100
BASE COUNT      207 a    257 c    179 g    163 t
ORIGIN

  Query Match             100.0%;  Score 806;  DB 6;  Length 806;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  806;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGAACGAGGGAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGAACGAGGGAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTG 60

Qy      61 TATCCTGCTTCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TATCCTGCTTCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGC 120

Qy     121 CTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAA 180

Qy     181 CCTGAAAGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCTGAAAGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCT 240

Qy     241 TTGGGTGGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TTGGGTGGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAG 300

Qy     301 CGGAGAACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CGGAGAACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCC 360

Qy     361 TTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAA 420

Qy     421 AGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAG 480

Qy     481 CCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACC 540

Qy     541 GAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGT 600

Qy     601 GCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAG 660

Qy     661 GTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTC 720

Qy     721 CTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAAT 780

Qy     781 AAATAAACTCAATGCAGACACAAAAA 806
          |||||||||||||||||||||||||||
Db     781 AAATAAACTCAATGCAGACACAAAAA 806


RESULT 2
AX464016
LOCUS       AX464016                804 bp    DNA     linear   PAT 16-JUL-2002
DEFINITION  Sequence 149 from Patent WO0140466.
ACCESSION   AX464016
VERSION     AX464016.1  GI:21899025
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,
            Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
            Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
            Wood,W.L. and Zhang,Z.
  TITLE     Secreted and transmembrane polypeptides and nucleic acids encoding
            same
  JOURNAL   Patent: WO 0140466-A 149 07-JUN-2001;
            Genentech Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .804
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      199 a    263 c    176 g    166 t
ORIGIN

  Query Match             97.0%;  Score 782;  DB 6;  Length 804;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy    610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy    670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy    730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy    790 CA 791
          ||
Db    803 CA 804
```

```
RESULT 3
AC022389
LOCUS       AC022389               198481 bp    DNA     linear   PRI 04-JUL-2002
DEFINITION  Homo sapiens chromosome 10 clone RP11-124L5, complete sequence.
ACCESSION   AC022389
VERSION     AC022389.9  GI:21321764
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2000) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   3  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   4  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUN-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   5  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   6  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   7  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Jun 4, 2002 this sequence version replaced gi:19071578.
FEATURES             Location/Qualifiers
     source          1. .198481
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-124L5"
                     /clone_lib="RPCI-11"
BASE COUNT    55275 a  44232 c  45110 g  53864 t
ORIGIN
```

```
  Query Match           76.4%;  Score 616;  DB 9;  Length 198481;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  616;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     187 AGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGT 246
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43093 AGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGT 43152

Qy     247 GGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGA 306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43153 GGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGA 43212

Qy     307 ACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACA 366
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43213 ACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACA 43272

Qy     367 GCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCC 426
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43273 GCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCC 43332

Qy     427 AGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACC 486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43333 AGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACC 43392

Qy     487 TCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCAT 546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43393 TCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCAT 43452

Qy     547 TCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGG 606
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43453 TCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGG 43512

Qy     607 TGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCAT 666
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43513 TGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCAT 43572

Qy     667 TCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCC 726
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43573 TCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCC 43632

Qy     727 TCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAA 786
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   43633 TCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAA 43692

Qy     787 ACTCAATGCAGACACA 802
           ||||||||||||||||
Db   43693 ACTCAATGCAGACACA 43708
```

```
RESULT 4
AX027767/c
LOCUS       AX027767                801 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 1 from Patent WO0043509.
ACCESSION   AX027767
VERSION     AX027767.1  GI:10188619
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   Vinals-Bassols,C.
  TITLE     Novel compounds
  JOURNAL   Patent: WO 0043509-A 1 27-JUL-2000;
            SMITHKLINE BEECHAM BIOLOG (BE) ; VINALS BASSOLS CARLOTA (BE)
FEATURES             Location/Qualifiers
     source          1. .801
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
BASE COUNT      165 a    181 c    255 g    200 t
ORIGIN

  Query Match           66.6%;  Score 537;  DB 6;  Length 801;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-312;
  Matches  537;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    261 TCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACCTCATGCCTGGC 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    537 TCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACCTCATGCCTGGC 478

Qy    321 ACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAA 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    477 ACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAA 418

Qy    381 TGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCC 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    417 TGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCC 358

Qy    441 TTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGC 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    357 TTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGC 298

Qy    501 CTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGG 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    297 CTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGG 238

Qy    561 TCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAA 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    237 TCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAA 178

Qy    621 GGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCA 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    177 GGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCA 118

Qy    681 TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTC 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    117 TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTC 58

Qy    741 TGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTCAATGCAG 797
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57 TGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTCAATGCAG 1


RESULT 5
AX027773/c
LOCUS       AX027773                797 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 7 from Patent WO0043509.
ACCESSION   AX027773
VERSION     AX027773.1  GI:10188625
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 797)
  AUTHORS   Vinals-Bassols,C.
  TITLE     Novel compounds
  JOURNAL   Patent: WO 0043509-A 7 27-JUL-2000;
            SMITHKLINE BEECHAM BIOLOG (BE) ; VINALS BASSOLS CARLOTA (BE)
FEATURES             Location/Qualifiers
     source          1. .797
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      164 a    180 c    254 g    199 t
ORIGIN

  Query Match           61.8%;  Score 498;  DB 6;  Length 797;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-289;
  Matches  498;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    261 TCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACCTCATGCCTGGC 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    533 TCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACCTCATGCCTGGC 474

Qy    321 ACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAA 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    473 ACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAA 414

Qy    381 TGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCC 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    413 TGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCC 354

Qy    441 TTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGC 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    353 TTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGC 294

Qy    501 CTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGG 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    293 CTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGG 234

Qy    561 TCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAA 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    233 TCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAA 174

Qy    621 GGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCA 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 GGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCA 114

Qy    681 TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTC 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    113 TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTC 54

Qy    741 TGCTGGAAAGAGCATCCA 758
          ||||||||||||||||||
Db     53 TGCTGGAAAGAGCATCCA 36


RESULT 6
AK025416
LOCUS       AK025416                2063 bp    mRNA    linear   PRI 29-SEP-2000
DEFINITION  Homo sapiens cDNA: FLJ21763 fis, clone COLF6967.
ACCESSION   AK025416
VERSION     AK025416.1  GI:10437924
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens colon mucosa cDNA to mRNA, clone_lib:ColF
            clone:COLF6967.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
            Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2063)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .2063
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /clone="COLF6967"
                     /tissue_type="colon mucosa"
                     /clone_lib="ColF"
                     /note="cloning vector pME18SFL3"
BASE COUNT       536 a    479 c    476 g    572 t
ORIGIN

  Query Match             59.6%;  Score 480;  DB 9;  Length 2063;
  Best Local Similarity   100.0%;  Pred. No. 4e-278;
  Matches  480;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 155

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     156 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 215

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     216 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 275

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     276 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 335

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     336 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 395

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     396 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 455

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     456 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 515

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     516 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 575


RESULT 7
AX351259
LOCUS       AX351259                485 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 6 from Patent WO0196390.
ACCESSION   AX351259
VERSION     AX351259.1  GI:18616606
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196390-A 6 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       109 a    155 c    124 g     97 t
ORIGIN

  Query Match             40.1%;  Score 323;  DB 6;  Length 485;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-183;
  Matches  323;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 119

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     120 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 179

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 239

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 299

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 359

Qy     310 TCATGCCTGGCACCTGAGGTACC 332
           |||||||||||||||||||||||
Db     360 TCATGCCTGGCACCTGAGGTACC 382


RESULT 8
AX340424/c
LOCUS       AX340424                382 bp    DNA     linear   PAT 10-JAN-2002
DEFINITION  Sequence 671 from Patent WO0196388.
ACCESSION   AX340424
VERSION     AX340424.1  GI:18136406
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Harlocker,S.L. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196388-A 671 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       76 a     94 c    122 g     86 t      4 others
ORIGIN

  Query Match             37.1%;  Score 299;  DB 6;  Length 382;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-168;
  Matches  299;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 261

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 201

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 141

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 81

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAAC 308
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAAC 22
```

```
RESULT 9
AX330090/c
LOCUS       AX330090                366 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 599 from Patent WO0194629.
ACCESSION   AX330090
VERSION     AX330090.1  GI:18103068
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 599 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .366
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       77 a     80 c    119 g     90 t
ORIGIN

  Query Match             24.3%;  Score 196;  DB 6;  Length 366;
  Best Local Similarity   99.6%;  Pred. No. 2.3e-106;
  Matches  246;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     357 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 307

Qy     417 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 247

Qy     477 TGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 536
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     246 TGAGCCAACCTCAACCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 187

Qy     537 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 596
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 127

Qy     597 GAGTGCA 603
           |||||||
Db     126 GAGTGCA 120


RESULT 10
AX335580/c
LOCUS       AX335580                366 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 6089 from Patent WO0194629.
ACCESSION   AX335580
VERSION     AX335580.1  GI:18126299
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 6089 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .366
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       77 a     80 c    119 g     90 t
ORIGIN

  Query Match             24.3%;  Score 196;  DB 6;  Length 366;
  Best Local Similarity   99.6%;  Pred. No. 2.3e-106;
  Matches  246;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     357 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 307

Qy     417 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 247

Qy     477 TGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 536
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     246 TGAGCCAACCTCAACCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 187

Qy     537 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 596
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 127

Qy     597 GAGTGCA 603
           |||||||
Db     126 GAGTGCA 120


RESULT 11
AX408291/c
LOCUS       AX408291                366 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 938 from Patent WO0229103.
ACCESSION   AX408291
VERSION     AX408291.1  GI:21440996
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 938 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1. .366
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. AA422086"
BASE COUNT       77 a     80 c    119 g     90 t
ORIGIN

  Query Match             24.3%;  Score 196;  DB 6;  Length 366;
  Best Local Similarity   99.6%;  Pred. No. 2.3e-106;
  Matches  246;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     357 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 307

Qy     417 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 247

Qy     477 TGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 536
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     246 TGAGCCAACCTCAACCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 187

Qy     537 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 596
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 127

Qy     597 GAGTGCA 603
           |||||||
```

```
Db    126 GAGTGCA 120


RESULT 12
AX342222
LOCUS       AX342222               4159 bp    DNA     linear   PAT 12-JAN-2002
DEFINITION  Sequence 8 from Patent WO0198497.
ACCESSION   AX342222
VERSION     AX342222.1  GI:18151765
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Polverino,A.J. and Luethy,R.
  TITLE     Secreted epithelial colon stromal-1 polypeptides, nucleic acids
            encoding the same and uses thereof
  JOURNAL   Patent: WO 0198497-A 8 27-DEC-2001;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4159
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     exon            1. .69
     exon            2627. .2725
     exon            4079. .4159
BASE COUNT     1009 a   1037 c    983 g   1110 t     20 others
ORIGIN

  Query Match             11.5%;  Score 93;  DB 6;  Length 4159;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-44;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     97 AGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCG 156
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2634 AGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCG 2693

Qy    157 AGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
          |||||||||||||||||||||||||||||||||
Db   2694 AGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 2726


RESULT 13
AL671190/c
LOCUS       AL671190               180176 bp    DNA     linear   ROD 24-MAY-2002
DEFINITION  Mouse DNA sequence from clone RP23-464L12 on chromosome 4, complete
            sequence.
ACCESSION   AL671190
VERSION     AL671190.9  GI:21214169
KEYWORDS    HTG.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 180176)
  AUTHORS   Griffiths,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 25, 2002 this sequence version replaced gi:21213389.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-464L12 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .180176
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="4"
                     /clone="RP23-464L12"
                     /clone_lib="RPCI-23"
BASE COUNT    49850 a   42493 c   40551 g   47282 t
ORIGIN

  Query Match             3.3%;  Score 27;  DB 10;  Length 180176;
  Best Local Similarity   100.0%;  Pred. No. 0.00046;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    757 CAAACAAACAAGTAATAAATAAATAAA 783
          |||||||||||||||||||||||||||
Db 148361 CAAACAAACAAGTAATAAATAAATAAA 148335


RESULT 14
AL357974/c
LOCUS       AL357974               93165 bp    DNA     linear   PRI 01-JUN-2001
DEFINITION  Human DNA sequence from clone RP11-398M15 on chromosome 1, complete
            sequence.
ACCESSION   AL357974
VERSION     AL357974.16  GI:14329558
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 93165)
  AUTHORS   Mashreghi-Mohammadi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 8, 2001 this sequence version replaced gi:13751416.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-398M15 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
```

```
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-398M15 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true right end of clone RP11-398M15 is at 93165 in this
            sequence. The true right end of clone RP11-110P20 is at 100 in this
            sequence.
FEATURES              Location/Qualifiers
     source           1. .93165
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /chromosome="1"
                      /clone="RP11-398M15"
                      /clone_lib="RPCI-11.2"
     repeat_region    843. .965
                      /note="L2 repeat: matches 2605. .2740 of consensus"
     repeat_region    1154. .1770
                      /note="L2 repeat: matches 1841. .2413 of consensus"
     repeat_region    1771. .2127
                      /note="MLT1A1 repeat: matches 1. .365 of consensus"
     repeat_region    2128. .2628
                      /note="L2 repeat: matches 1216. .1841 of consensus"
     repeat_region    3949. .3990
                      /note="LTR40b repeat: matches 421. .462 of consensus"
     repeat_region    7352. .7720
                      /note="L1PA4 repeat: matches 5777. .6146 of consensus"
     repeat_region    8235. .8278
                      /note="11 copies 4 mer caca 100% conserved"
     repeat_region    8831. .8980
                      /note="75 copies 2 mer at 85% conserved"
     repeat_region    8837. .8980
                      /note="36 copies 4 mer atat 86% conserved"
     repeat_region    9022. .9093
                      /note="L2 repeat: matches 2679. .2750 of consensus"
     repeat_region    9284. .9367
                      /note="L1M4 repeat: matches 3167. .3240 of consensus"
     repeat_region    9368. .10098
                      /note="L1PA13 repeat: matches 5419. .6155 of consensus"
     repeat_region    10099. .10174
                      /note="L1M4 repeat: matches 3061. .3167 of consensus"
     repeat_region    10237. .12387
                      /note="L1MEc repeat: matches 1379. .3033 of consensus"
     repeat_region    12458. .13627
                      /note="L1MEc repeat: matches 102. .1288 of consensus"
     repeat_region    14898. .21039
                      /note="L1PA3 repeat: matches 1. .6145 of consensus"
     repeat_region    22297. .23011
                      /note="LTR1 repeat: matches 1. .718 of consensus"
     repeat_region    23284. .23408
                      /note="MIR repeat: matches 118. .248 of consensus"
     repeat_region    23442. .24061
                      /note="MER67C repeat: matches 1. .709 of consensus"
     repeat_region    24234. .24297
                      /note="MER92C repeat: matches 489. .552 of consensus"
     repeat_region    24564. .25039
                      /note="MER89 repeat: matches 1. .559 of consensus"
     repeat_region    25623. .25911
                      /note="AluSc repeat: matches 1. .301 of consensus"
     repeat_region    28551. .30612
                      /note="L1MC2 repeat: matches 4265. .6328 of consensus"
     repeat_region    31945. .32066
                      /note="MSTB repeat: matches 305. .426 of consensus"
     repeat_region    32574. .33996
                      /note="L1MEc repeat: matches 1317. .2374 of consensus"
     repeat_region    33997. .34097
                      /note="L2 repeat: matches 2146. .2247 of consensus"
     repeat_region    34152. .34322
                      /note="L2 repeat: matches 2552. .2732 of consensus"
     repeat_region    35014. .35066
                      /note="L2 repeat: matches 2657. .2709 of consensus"
     repeat_region    38011. .38055
                      /note="5S repeat: matches 1. .45 of consensus"
     repeat_region    38062. .38665
                      /note="L1PA11 repeat: matches 5549. .6165 of consensus"
     repeat_region    40064. .40625
                      /note="HERVL repeat: matches 4356. .4930 of consensus"
     repeat_region    40716. .40783
                      /note="17 copies 4 mer ggag 69% conserved"
     repeat_region    41270. .41408
                      /note="L2 repeat: matches 1273. .1413 of consensus"
     repeat_region    41409. .41691
                      /note="AluY repeat: matches 1. .309 of consensus"
     repeat_region    41692. .41765
                      /note="L2 repeat: matches 1202. .1273 of consensus"
     repeat_region    41766. .42192
                      /note="MSTD repeat: matches 1. .394 of consensus"
     repeat_region    42193. .42252
                      /note="L2 repeat: matches 1143. .1202 of consensus"
     repeat_region    43672. .43854
                      /note="MER5A repeat: matches 1. .183 of consensus"
     repeat_region    44566. .44647
                      /note="L1ME repeat: matches 5395. .5473 of consensus"
     repeat_region    45074. .45317
                      /note="MIR repeat: matches 4. .259 of consensus"
     repeat_region    45463. .45656
                      /note="MIR repeat: matches 74. .262 of consensus"
     repeat_region    45929. .46050
                      /note="L2 repeat: matches 2626..2750 of consensus"
     repeat_region    46167. .46272
                      /note="L2 repeat: matches 2384. .2501 of consensus"
     repeat_region    47122. .47163
                      /note="21 copies 2 mer ac 92% conserved"
     repeat_region    47124. .47163
                      /note="10 copies 4 mer acac 95% conserved"
     repeat_region    47825. .48134
                      /note="AluSg repeat: matches 1. .310 of consensus"
     repeat_region    48197. .48292
                      /note="L2 repeat: matches 2662. .2739 of consensus"
     repeat_region    48293. .48670
                      /note="THE1C repeat: matches 1. .371 of consensus"
     repeat_region    48671. .48763
                      /note="L2 repeat: matches 2564. .2662 of consensus"
     repeat_region    48766. .48933
                      /note="MER69A repeat: matches 5. .177 of consensus"
     repeat_region    49075. .49393
                      /note="L2 repeat: matches 69. .386 of consensus"
     repeat_region    51782. .52105
                      /note="MER44A repeat: matches 3. .333 of consensus"
     repeat_region    53670. .54781
                      /note="L1PA15 repeat: matches 5065. .6157 of consensus"
     repeat_region    54782. .55089
                      /note="AluSp repeat: matches 3. .310 of consensus"
     repeat_region    55090. .55691
                      /note="L1PA15 repeat: matches 4480. .5065 of consensus"
     repeat_region    55692. .56707
                      /note="L1P repeat: matches 2855. .3886 of consensus"
     repeat_region    56880. .56986
                      /note="L1M4 repeat: matches 2949. .3054 of consensus"
     repeat_region    57122. .57177
                      /note="14 copies 4 mer tata 85% conserved"
     repeat_region    57743. .57880
                      /note="L2 repeat: matches 2621. .2742 of consensus"
     repeat_region    58525. .58570
                      /note="MSTA repeat: matches 381. .426 of consensus"
     repeat_region    58573. .58866
                      /note="AluSp repeat: matches 4. .307 of consensus"
     repeat_region    58867. .58942
                      /note="19 copies 4 mer ggaa 89% conserved"
     repeat_region    58962. .59333
                      /note="MSTA repeat: matches 6. .399 of consensus"
     repeat_region    59428. .59494
                      /note="L2 repeat: matches 1951. .2023 of consensus"
     repeat_region    59495. .59717
                      /note="MER58A repeat: matches 2. .223 of consensus"
     repeat_region    59718. .59844
```

```
                     /note="L2 repeat: matches 1849. .1951 of consensus"
     repeat_region   61894. .62182
                     /note="AluY repeat: matches 1. .289 of consensus"
     repeat_region   62183. .62230
                     /note="12 copies 4 mer taaa 93% conserved"
     repeat_region   62351. .62659
                     /note="AluSp repeat: matches 1. .311 of consensus"
     repeat_region   62847. .63179
                     /note="MLT1A1 repeat: matches 1. .365 of consensus"
     repeat_region   64094. .64321
                     /note="L1M4c repeat: matches 1515. .1742 of consensus"
     repeat_region   64649. .64953
                     /note="MLT1E repeat: matches 264. .568 of consensus"

  Query Match             3.1%;  Score 25;  DB 9;  Length 93165;
  Best Local Similarity   100.0%;  Pred. No. 0.0073;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     763 AACAAGTAATAAATAAATAAATAAA 787
           |||||||||||||||||||||||||
Db   55365 AACAAGTAATAAATAAATAAATAAA 55341


RESULT 15
AC022829/c
LOCUS       AC022829              152245 bp    DNA     linear   HTG 16-MAR-2000
DEFINITION  Homo sapiens clone RP11-121C1, WORKING DRAFT SEQUENCE, 20 unordered
            pieces.
ACCESSION   AC022829
VERSION     AC022829.3  GI:7249268
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152245)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-121C1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152245)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 16, 2000 this sequence version replaced gi:6984468.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L6137
                Center clone name: 121_C_1
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 142478 bases at least Q40
                Consensus quality: 147185 bases at least Q30
                Consensus quality: 148983 bases at least Q20
                Insert size: 150000; agarose-fp
                Insert size: 150345; sum-of-contigs
                Quality coverage: 4.3 in Q20 bases; agarose-fp
                Quality coverage: 4.3 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 20 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1283: contig of 1283 bp in length
            *     1284 1383: gap of      100 bp
            *     1384     1916: contig of 533 bp in length
            *     1917 2016: gap of      100 bp
            *     2017     3647: contig of 1631 bp in length
            *     3648 3747: gap of      100 bp
            *     3748     5481: contig of 1734 bp in length
            *     5482 5581: gap of      100 bp
            *     5582     6718: contig of 1137 bp in length
            *     6719 6818: gap of      100 bp
            *     6819    10904: contig of 4086 bp in length
            *    10905 11004: gap of      100 bp
            *    11005    14393: contig of 3389 bp in length
            *    14394 14493: gap of      100 bp
            *    14494    18255: contig of 3762 bp in length
            *    18256 18355: gap of      100 bp
            *    18356    23193: contig of 4838 bp in length
            *    23194 23293: gap of      100 bp
            *    23294    28506: contig of 5213 bp in length
            *    28507 28606: gap of      100 bp
            *    28607    36308: contig of 7702 bp in length
            *    36309 36408: gap of      100 bp
            *    36409    43542: contig of 7134 bp in length
            *    43543 43642: gap of      100 bp
            *    43643    51047: contig of 7405 bp in length
            *    51048 51147: gap of      100 bp
            *    51148    61019: contig of 9872 bp in length
            *    61020 61119: gap of      100 bp
            *    61120    74018: contig of 12899 bp in length
            *    74019 74118: gap of      100 bp
            *    74119    84938: contig of 10820 bp in length
            *    84939 85038: gap of      100 bp
            *    85039   100034: contig of 14996 bp in length
            *   100035 100134: gap of      100 bp
            *   100135   112980: contig of 12846 bp in length
            *   112981 113080: gap of      100 bp
            *   113081   130270: contig of 17190 bp in length
            *   130271 130370: gap of      100 bp
            *   130371   152245: contig of 21875 bp in length.
FEATURES             Location/Qualifiers
     source          1. .152245
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-121C1"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1283
                     /note="assembly_fragment"
     misc_feature    1384. .1916
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     misc_feature    2017. .3647
                     /note="assembly_fragment"
     misc_feature    3748. .5481
```

```
                     /note="assembly_fragment"
     misc_feature    5582. .6718
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    6819. .10904
                     /note="assembly_fragment"
     misc_feature    11005. .14393
                     /note="assembly_fragment"
     misc_feature    14494. .18255
                     /note="assembly_fragment"
     misc_feature    18356. .23193
                     /note="assembly_fragment"
     misc_feature    23294. .28506
                     /note="assembly_fragment"
     misc_feature    28607. .36308
                     /note="assembly_fragment"
     misc_feature    36409. .43542
                     /note="assembly_fragment"
     misc_feature    43643. .51047
                     /note="assembly_fragment"
     misc_feature    51148. .61019
                     /note="assembly_fragment"
     misc_feature    61120. .74018
                     /note="assembly_fragment"
     misc_feature    74119. .84938
                     /note="assembly_fragment"
     misc_feature    85039. .100034
                     /note="assembly_fragment"
     misc_feature    100135. .112980
                     /note="assembly_fragment"
     misc_feature    113081. .130270
                     /note="assembly_fragment"
     misc_feature    130371. .152245
                     /note="assembly_fragment"
BASE COUNT    47920 a  28046 c  27243 g  47134 t   1902 others
ORIGIN

  Query Match             3.1%;  Score 25;  DB 2;  Length 152245;
  Best Local Similarity   100.0%;  Pred. No. 0.0075;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     770 AATAAATAAATAAATAAACTCAATG 794
           |||||||||||||||||||||||||
Db   43991 AATAAATAAATAAATAAACTCAATG 43967


Search completed: May 17, 2003, 16:50:03
Job time : 3076 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 11:17:06 ; Search time 267 Seconds
(without alignments)
6798.167 Million cell updates/sec

Title: US-09-724-000A-4
Perfect score: 806
Sequence: 1 ggaacgagggaaaatctgcc.....actcaatgcagacacaaaaa 806

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match %	Length	DB	ID	Description
	1	806	100.0	806	24	AAD27025	Human Secreted epi
	2	782	97.0	804	22	AAS21318	Human cDNA sequenc
	3	732	90.8	908	22	AAF97898	Human secreted pro
c	4	537	66.6	801	21	AAA72224	Human CASB611 cDNA
c	5	498	61.8	797	21	AAA72230	Human CASB gene pa
	6	323	40.1	485	24	ABK27569	Human colon cancer
c	7	299	37.1	382	24	ABL37082	Human colon tumour
	8	231	28.7	487	22	AAS39569	cDNA encoding nove
	9	231	28.7	487	22	AAK88908	Human digestive sy
	10	220	27.3	227	22	AAS02026	Bladder cancer-ass
c	11	196	24.3	366	24	ABN94440	Gene #938 used to
c	12	196	24.3	366	24	ABL62262	Colon adenocarcino
c	13	196	24.3	366	24	ABL67752	Oesophagus cancer
c	14	133	16.5	629	24	ABQ57070	Human colon cancer
	15	93	11.5	4159	24	AAD27026	Human Secreted epi
	16	60	7.4	60	24	ABN43987	Human spliced tran
c	17	23	2.9	151	21	AAC20052	Human secreted pro
c	18	23	2.9	7355	24	ABN80297	Human chemically m
c	19	23	2.9	8911	22	ABA06809	Human genomic DNA
c	20	23	2.9	8911	22	AAS41748	Genomic sequence #
c	21	23	2.9	119198	24	ABK83565	Human cDNA differe
c	22	22	2.7	344	21	AAC23854	Human secreted pro
	23	22	2.7	377	24	ABK44861	cDNA encoding colo
	24	22	2.7	377	24	ABK45072	cDNA encoding colo
c	25	22	2.7	494	24	ABK27657	Human colon cancer
c	26	22	2.7	801	22	AAI97646	Human neuroblastom
c	27	22	2.7	1844	21	AAF16260	Human prostate can
c	28	22	2.7	2368	23	ABL28260	Drosophila melanog
c	29	22	2.7	13322	22	AAL06979	Human reproductive
c	30	22	2.7	143306	24	ABK49586	Human transporter
	31	21	2.6	21	24	AAD27027	Human Secreted epi
	32	21	2.6	1693	22	AAF63803	Human secreted pro
	33	21	2.6	5491	22	AAK69044	Human immune/haema
c	34	21	2.6	5586	24	ABL32648	Human immune syste
c	35	21	2.6	7669	24	ABL32628	Human immune syste
c	36	21	2.6	8101	24	ABL33979	Human immune syste
c	37	21	2.6	9299	24	ABL33724	Human immune syste
c	38	21	2.6	15500	22	AAS46507	Tumour suppressor
c	39	21	2.6	125439	24	ABQ88177	Human osteoblast d
c	40	20	2.5	186	21	AAC27822	Human secreted pro
c	41	20	2.5	441	21	AAC28871	Human secreted pro
	42	20	2.5	442	22	AAH33689	Human colon cancer
	43	20	2.5	452	17	AAT42809	Polymorphic locus
c	44	20	2.5	463	22	AAK90944	Human digestive sy
c	45	20	2.5	463	22	AAS31979	Human liver associ

ALIGNMENTS

RESULT 1
AAD27025
ID AAD27025 standard; DNA; 806 BP.
XX
AC AAD27025;
XX
DT 09-APR-2002 (first entry)
XX
DE Human Secreted epithelial colon stromal-1 (Secs-1) DNA.
XX
KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW human; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 29..274
FT /*tag= a
FT /product= "Human Secs-1 protein"
FT sig_peptide 29..100
FT /*tag= b
FT mat_peptide 101..271
FT /*tag= c
FT /product= "Mature human Secs-1 protein"
XX
PN WO200198497-A1.
XX

PD 27-DEC-2001.
XX
PF 28-NOV-2000; 2000WO-US32479.
XX
PR 21-JUN-2000; 2000US-0599087.
PR 28-NOV-2000; 2000US-0724000.
XX
PA (AMGE-) AMGEN INC.
XX
PI Polverino AJ, Luethy R;
XX
DR WPI; 2002-122281/16.
DR P-PSDB; AAE16481.
XX
PT Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT useful for diagnosing, treating and preventing hematopoietic disorder,
PT osteoporosis, Paget's disease, cancer, diabetes -
XX
PS Claim 1; Fig 2; 134pp; English.
XX
CC The present invention relates to an isolated murine or human secreted
CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC therapy and cell therapy. Secs-1 is useful for identifying a compound
CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC preventing or ameliorating a disease condition such as haematopoietic
CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC is also useful for diagnosing a pathological condition which involves
CC determining the presence or amount of Secs-1 or polypeptide encoded by
CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC susceptibility to pathological condition based on the presence or amount
CC of expression of the polypeptide. The present sequence is human Secs-1
CC DNA.
XX
SQ Sequence 806 BP; 207 A; 257 C; 179 G; 163 T; 0 other;

Query Match 100.0%; Score 806; DB 24; Length 806;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGAACGAGGGAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGAACGAGGGAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTG 60

Qy      61 TATCCTGCTTCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TATCCTGCTTCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGC 120

Qy     121 CTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAA 180

Qy     181 CCTGAAAGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCTGAAAGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCT 240

Qy     241 TTGGGTGGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TTGGGTGGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAG 300

Qy     301 CGGAGAACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CGGAGAACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCC 360

Qy     361 TTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAA 420

Qy     421 AGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAG 480

Qy     481 CCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACC 540

Qy     541 GAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGT 600

Qy     601 GCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCAGGGTGGCAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAG 660

Qy     661 GTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTCCATTCAGCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTC 720

Qy     721 CTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CTAGCCTCCTGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAAT 780

Qy     781 AAATAAACTCAATGCAGACACAAAAA 806
          |||||||||||||||||||||||||||
Db     781 AAATAAACTCAATGCAGACACAAAAA 806
```

RESULT 2
AAS21318
ID AAS21318 standard; cDNA; 804 BP.
XX
AC AAS21318;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human cDNA sequence encoding for PRO3446 polypeptide.
XX
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200140466-A2.
XX
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000WO-US32678.
XX
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.

```
PR   30-MAY-2000; 2000WO-US14941.
PR   02-JUN-2000; 2000WO-US15264.
PR   10-NOV-2000; 2000WO-US30873.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI   Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI   Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR   WPI; 2001-408281/43.
DR   P-PSDB; AAU12246.
XX
PT   Isolated, secretory and transmembrane PRO polypeptide used to detect
PT   other PRO polypeptides, link bioactive molecules to cells expressing
PT   PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT   lung, breast, prostate, cervical -
XX
PS   Claim 3; Fig 149; 813pp; English.
XX
CC   AAS21244-AAS21518 encode for novel human secretory and transmembrane
CC   PRO polypeptides. The PRO polypeptides are useful to detect other
CC   PRO polypeptides, to link bioactive molecules to cells expressing
CC   PRO polypeptides, to modulate biological activities of cells expressing
CC   PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC   breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC   polypeptide expression in a cell sample to that in a control sample.
CC   Some of the 275 sequences are also useful to stimulate the release of
CC   tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC   proliferation or differentiation of chondrocytes, the proliferation or
CC   gene expression in pericyte cells, the release of proteoglycans from
CC   cartilage, the proliferation of inner ear utricular supporting cells or
CC   of T-lymphocytes, the release of a cytokine from peripheral blood
CC   monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC   the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC   skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC   to factor VIIA. The PRO polypeptides can be used in assays to identify
CC   molecules involved in binding interactions. The polynucleotides encoding
CC   PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC   transgenic or knock out animals and can be used in gene therapy.
XX
SQ   Sequence 804 BP; 199 A; 263 C; 176 G; 166 T; 0 other;

  Query Match            97.0%;  Score 782;  DB 22;  Length 804;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 82

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 142

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 202

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 262

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 322

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 382

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 442

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 502

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 562

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 742

Qy     730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 802

Qy     790 CA 791
           ||
Db     803 CA 804


RESULT 3
AAF97898
ID   AAF97898 standard; cDNA; 908 BP.
XX
AC   AAF97898;
XX
DT   01-JUN-2001  (first entry)
XX
DE   Human secreted protein cDNA, SEQ ID NO: 25.
XX
KW   Human; secreted protein; immunomodulatory; antisclerotic;
KW   dermatological; antiinflammatory; anti-HIV; cytostatic; cardiant;
KW   vascular; anti-angiogenic; ophthalmological; neuroprotectant;
KW   nootropic; anticonvulsant; antialzheimers; antiparkinsonian;
KW   antimicrobial; vulnerary; vaccine; gene therapy; cancer;
KW   protein coordinate data; infection; ss.
XX
OS   Homo sapiens.
XX
PN   WO200121658-A1.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-US26013.
XX
PR   24-SEP-1999;   99US-0155709.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
PI   Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
PI   Young PE, Wei P, Florence KA;
XX
DR   WPI; 2001-235311/24.
XX
PT   Nucleic acids encoding 32 human secreted polypeptides, useful for
PT   preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT   disease and diabetic retinopathy -
XX
PS   Claim 1; Page 730; 890pp; English.
XX
CC   The present sequence encodes one of 32 novel human secreted polypeptides.
```

```
CC   The nucleic acid molecules and polypeptides they encode may be used in
CC   the prevention, diagnosis and treatment of diseases such as
CC   immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus
CC   and human immuno-deficiency virus (HIV) infections), hyperproliferative
CC   disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
CC   (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary
CC   arteriosclerosis), angiogenic disorders (e.g. corneal graft
CC   neovascularisation and diabetic retinopathy), neurological disorders
CC   (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
CC   infectious diseases and/or for promoting wound healing, regeneration
CC   and/or chemotaxis. The nucleic acid molecules may be used to produce the
CC   secreted polypeptides. They may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acid
CC   sequences in samples. The polypeptides may be used as antigens in the
CC   production of antibodies and in assays to identify modulators of
CC   their expression and activity.
XX
SQ   Sequence 908 BP; 242 A; 290 C; 197 G; 175 T; 4 other;

  Query Match             90.8%;  Score 732;  DB 22;  Length 908;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 125

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 185

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 245

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     246 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 305

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 365

Qy     310 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 TCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCA 425

Qy     370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     426 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGA 485

Qy     430 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     486 ACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCA 545

Qy     490 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     546 CCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCC 605

Qy     550 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 609
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     606 ACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGWGG 665

Qy     610 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     666 CAAGCACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 725

Qy     670 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     726 GCCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCC 785

Qy     730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     786 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACT 845

Qy     790 CAA 792
           |||
Db     846 CAA 848


RESULT 4
AAA72224/c
ID   AAA72224 standard; cDNA; 801 BP.
XX
AC   AAA72224;
XX
DT   06-DEC-2000  (first entry)
XX
DE   Human CASB611 cDNA.
XX
KW   Human; CASB611; colon-specific expression; expressed sequence tag;
KW   EST; colon cancer; tumour; autoimmune disease; diagnosis;
KW   disease susceptilibity; prophylaxis; genetic vaccine; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200043509-A2.
XX
PD   27-JUL-2000.
XX
PF   17-JAN-2000; 2000WO-EP00346.
XX
PR   19-JAN-1999;   99GB-0001078.
PR   29-JAN-1999;   99GB-0002090.
PR   01-FEB-1999;   99GB-0002163.
PR   01-FEB-1999;   99GB-0002168.
PR   01-FEB-1999;   99GB-0002169.
PR   07-APR-1999;   99GB-0007901.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Vinals-Bassols C;
XX
DR   WPI; 2000-482912/42.
XX
PT   New isolated polynucleotide useful for diagnosis and/or treatment of
PT   colon cancer and autoimmune disease -
XX
PS   Claim 3; Page 34; 41pp; English.
XX
CC   This sequence represents human CASB611 cDNA. This gene exhibits
CC   colon-specific expression and is highly expressed in the rectum.
CC   The invention relates to human CASB cDNA sequences CASB611, CASB500,
CC   CASB501, CASB502, CASB505 and CASB507 (AAA72224-A72229, respectively)
CC   and also to human CASB partial cDNA sequences (AAA72230-A72235) derived
CC   from expressed sequence tags (ESTs). Expression of the human CASB genes
CC   (with the exception of CASB611) is associated with colon tumours, and
CC   the encoded proteins (sequences not given in the specification)
CC   represent colon tumour-associated antigens. The cDNA sequences may be
CC   used in diagnosing the presence or a susceptibility to a disease related
CC   to the presence, expression or activity of CASB genes. Such diseases
CC   include autoimmune diseases and especially colon cancer. The nucleic
CC   acid sequences may also be used in genetic vaccines for the prophylaxis
CC   or therapeutic treatment of colon cancer and autoimmune diseases.
XX
SQ   Sequence 801 BP; 165 A; 181 C; 255 G; 200 T; 0 other;

  Query Match             66.6%;  Score 537;  DB 21;  Length 801;
  Best Local Similarity   100.0%;  Pred. No. 2e-246;
  Matches  537;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     261 TCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACCTCATGCCTGGC 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     537 TCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACCTCATGCCTGGC 478

Qy     321 ACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAA 380
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    477 ACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAA 418

Qy    381 TGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCC 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    417 TGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCC 358

Qy    441 TTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGC 500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    357 TTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGC 298

Qy    501 CTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGG 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    297 CTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGG 238

Qy    561 TCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAA 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    237 TCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAA 178

Qy    621 GGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCA 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    177 GGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCA 118

Qy    681 TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTC 740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    117 TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTC 58

Qy    741 TGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTCAATGCAG 797
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57 TGCTGGAAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTCAATGCAG 1
```

RESULT 5
AAA72230/c
ID AAA72230 standard; cDNA; 797 BP.
XX
AC AAA72230;
XX
DT 06-DEC-2000 (first entry)
XX
DE Human CASB gene partial cDNA, SEQ ID NO:7.
XX
KW Human; CASB gene; overexpression; colon tumour-associated antigen;
KW expressed sequence tag; EST; colon cancer; tumour; autoimmune disease;
KW diagnosis; disease susceptilibity; prophylaxis; genetic vaccine;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200043509-A2.
XX
PD 27-JUL-2000.
XX
PF 17-JAN-2000; 2000WO-EP00346.
XX
PR 19-JAN-1999; 99GB-0001078.
PR 29-JAN-1999; 99GB-0002090.
PR 01-FEB-1999; 99GB-0002163.
PR 01-FEB-1999; 99GB-0002168.
PR 01-FEB-1999; 99GB-0002169.
PR 07-APR-1999; 99GB-0007901.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Vinals-Bassols C;
XX
DR WPI; 2000-482912/42.
XX
PT New isolated polynucleotide useful for diagnosis and/or treatment of
PT colon cancer and autoimmune disease -
XX
PS Disclosure; Page 35-36; 41pp; English.
XX
CC Sequences AAA72230-A72235 represent human CASB gene partial cDNA
CC sequences which are derived from expressed sequence tags (ESTs). The
CC invention relates to human CASB cDNA sequences CASB611, CASB500,
CC CASB501, CASB502, CASB505 and CASB507 (AAA72224-A72229, respectively) and
CC also to these human CASB partial cDNA sequences. Expression of the human
CC CASB genes (with the exception of CASB611) is associated with colon
CC tumours, and the encoded proteins (sequences not given in the
CC specification) represent colon tumour-associated antigens. The cDNA
CC sequences may be used in diagnosing the presence or a susceptibility to
CC a disease related to the presence, expression or activity of CASB genes.
CC Such diseases include autoimmune diseases and especially colon cancer.
CC The nucleic acid sequences may also be used in genetic vaccines for the
CC prophylaxis or therapeutic treatment of colon cancer and autoimmune
CC diseases.
XX
SQ Sequence 797 BP; 164 A; 180 C; 254 G; 199 T; 0 other;

  Query Match   61.8%; Score 498; DB 21; Length 797;
  Best Local Similarity 100.0%; Pred. No. 8.5e-228;
  Matches 498; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    261 TCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACCTCATGCCTGGC 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    533 TCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACCTCATGCCTGGC 474

Qy    321 ACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAA 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    473 ACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCACAGCAGTGAGCTGCAA 414

Qy    381 TGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCC 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    413 TGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCCACGTCC 354

Qy    441 TTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGC 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    353 TTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCACAGGGC 294

Qy    501 CTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGG 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    293 CTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCATGACCGG 234

Qy    561 TCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAA 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    233 TCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAA 174

Qy    621 GGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCA 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 GGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCA 114

Qy    681 TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTC 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    113 TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTC 54

Qy    741 TGCTGGAAAGAGCATCCA 758
          ||||||||||||||||||
Db     53 TGCTGGAAAGAGCATCCA 36
```

RESULT 6
ABK27569
ID ABK27569 standard; cDNA; 485 BP.
XX
AC ABK27569;
XX
DT 09-APR-2002 (first entry)
XX
DE Human colon cancer expressed sequence tag, Seq ID no 6.
XX
KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
KW expressed sequence tag.
XX
OS Homo sapiens.
XX

PN WO200196390-A2.
XX
PD 20-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US18577.
XX
PR 09-JUN-2000; 2000US-210821P.
PR 18-DEC-2000; 2000US-256571P.
PR 10-MAY-2001; 2001US-290240P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;
XX
DR WPI; 2002-139708/18.
XX
PT Novel isolated polynucleotide encoding a polypeptide comprising a
PT portion of colon tumour protein, useful for detection, diagnosis and
PT therapy of human colon cancer -
XX
PS Claim 1; Page 151-152; 220pp; English.
XX
CC The invention relates to an isolated polynucleotide (I) encoding a
CC polypeptide (II) comprising at least a portion of a colon tumour
CC protein. (I), (II) and antibody (III) to (II) are useful for determining
CC the presence of a cancer in a patient. (I), (II) or antigen presenting
CC cells expressing (I) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein, by contacting T cells with (I), (II) or
CC antigen-presenting cells that express (I), under conditions and for a
CC time sufficient to permit the stimulation and/or expansion of T cells.
CC (I), (II), or antigen presenting cells that express (II) are useful for
CC treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
CC isolated from a patient with (I), (II) or antigen presenting cells that
CC express (II), such that T cells proliferate, and administering to the
CC patient an effective amount of the proliferated T cells, thus inhibiting
CC the development of a cancer in the patient. (I) or (II) is useful in
CC vaccines and pharmaceutical compositions for prevention and treatment
CC of colon malignancies and for the diagnosis and monitoring of such
CC cancers. (I), (II) or (III) is useful for detection, diagnosis and/or
CC therapy of human colon cancer. (I) is useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of (II) in tumour cells.
CC ABK27564-ABK27807 represent novel human colon cancer coding
CC sequences and primers of the invention.
XX
SQ Sequence 485 BP; 109 A; 155 C; 124 G; 97 T; 0 other;

Query Match 40.1%; Score 323; DB 24; Length 485;
Best Local Similarity 100.0%; Pred. No. 3.3e-144;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 119

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     120 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 179

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 239

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 299

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAACC 359

Qy     310 TCATGCCTGGCACCTGAGGTACC 332
           |||||||||||||||||||||||
```

Db 360 TCATGCCTGGCACCTGAGGTACC 382

RESULT 7
ABL37082/c
ID ABL37082 standard; cDNA; 382 BP.
XX
AC ABL37082;
XX
DT 08-APR-2002 (first entry)
XX
DE Human colon tumour antigen polynucleotide SEQ ID NO:671.
XX
KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200196388-A2.
XX
PD 20-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US18557.
XX
PR 09-JUN-2000; 2000US-210899P.
PR 20-FEB-2001; 2001US-270216P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Harlocker SL, Secrist H;
XX
DR WPI; 2002-114514/15.
XX
PT Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient -
XX
PS Claim 1; SEQ ID 671; 105pp; English.
XX
CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC production. (I) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (I) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (I)
CC can be used in the diagnosis of a colon tumour.
XX
SQ Sequence 382 BP; 76 A; 94 C; 122 G; 86 T; 4 other;

Query Match 37.1%; Score 299; DB 24; Length 382;
Best Local Similarity 100.0%; Pred. No. 9.7e-133;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      10 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 GAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCT 261

Qy      70 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 TCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGG 201

Qy     130 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 CAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 141

Qy     190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGT 81

Qy     250 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAAC 308
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     80 GCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCGGAGAAC 22

RESULT 8
AAS39569
ID   AAS39569 standard; cDNA; 487 BP.
XX
AC   AAS39569;
XX
DT   17-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human colon associated polypeptide #222.
XX
KW   Human; colon cancer; congenital abnormality; infection; colitis;
KW   inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KW   intestinal inflammatory disorder; malabsorption syndrome; gastric;
KW   sigmoid disease; antibacterial; antiviral; antiinflammatory;
KW   cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155302-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01240.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
```

```
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465567/50.
DR   P-PSDB; AAU22689.
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   disorders related to the colon including colon cancers and also for
PT   testing and detection e.g. diagnosis -
XX
PS   Claim 4; SEQ ID No 232; 562pp; English.
XX
CC   The present invention relates to the isolation of novel human colon
CC   associated polypeptides (AAU22468-AAU22701), and the cDNA and genomic
CC   sequences encoding for them. The sequences of the invention are useful
CC   in the diagnosis, treatment, prevention and/or prognosis of disorders
CC   of the colon including colon cancer, congenital abnormalities
CC   (e.g. atresia and stenosis), bacterial and viral infections,
CC   inflammatory bowel disease (IBD), neoplastic cell disorders
CC   (e.g. polyps and adenomas, intestinal inflammatory disorders, colitis,
CC   colonic inflammation, diarrhoea and dysentery, malabsorption syndromes
CC   (e.g. lactose intolerance), intestinal obstruction and sigmoid diseases.
CC   The polynucleotides sequences of the invention can also be used in gene
CC   therapy. AAS39348-AAS39581 represent cDNA sequences encoding for the
CC   novel human colon associated polypeptides of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 487 BP; 135 A; 156 C; 100 G; 95 T; 1 other;

  Query Match             28.7%;  Score 231;  DB 22;  Length 487;
  Best Local Similarity   99.6%;  Pred. No. 2.9e-100;
  Matches  281;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     375 CTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCC 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69 CTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCC 128

Qy     435 ACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCA 494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129 ACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCA 188

Qy     495 CAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCAT 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189 CAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCAT 248

Qy     555 GACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGC 614
           ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     249 GACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCATGGTGGCAAGC 308

Qy     615 ACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCT 656
           |||||||||||||||||||||||||||||||||||||||||||
Db     309 ACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCT 350

RESULT 9
AAK88908
ID   AAK88908 standard; cDNA; 487 BP.
XX
AC   AAK88908;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human digestive system antigen coding sequence SEQ ID NO: 1224.
XX
KW   Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW   ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW   digestive system disorder; Meckel's diverticulum; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155314-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01324.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
```

PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-502630/55.
DR P-PSDB; AAM93135.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognozing disorders of the
PT digestive system, particularly cancer and cancer metastases -
XX
PS Claim 1; SEQ ID NO 1224; 986pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a cDNA encoding a digestive
CC system antigen of the invention.
XX
SQ Sequence 487 BP; 135 A; 156 C; 100 G; 95 T; 1 other;

Query Match 28.7%; Score 231; DB 22; Length 487;
Best Local Similarity 99.6%; Pred. No. 2.9e-100;
Matches 281; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    375 CTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCC 434
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 CTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAGAACTCC 128

Qy    435 ACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCA 494
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 ACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACCCCA 188

Qy    495 CAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCAT 554
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 CAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCCACCAT 248

Qy    555 GACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGC 614
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    249 GACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCATGGTGGCAAGC 308

Qy    615 ACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCT 656
          |||||||||||||||||||||||||||||||||||||||||
Db    309 ACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCT 350
```

RESULT 10
AAS02026
ID AAS02026 standard; cDNA; 227 BP.
XX
AC AAS02026;
XX
DT 16-JUL-2001 (first entry)
XX
DE Bladder cancer-associated sequence, TCC96C7.
XX

```
KW   Bladder; cancer; transitional cell carcinoma; diagnostic; TCC96C7; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   polyA_site      218
FT                   /*tag=  a
XX
PN   WO200122864-A2.
XX
PD   05-APR-2001.
XX
PF   27-SEP-2000; 2000WO-US41005.
XX
PR   27-SEP-1999;   99US-0156153.
XX
PA   (QUAR-) QUARK BIOTECH INC.
PA   (KOHN/) KOHN K I.
XX
PI   Feinstein E,  Mor O;
XX
DR   WPI; 2001-258076/26.
XX
PT   New polynucleotide sequences upregulated in bladder cancer for
PT   diagnosing bladder cancer and inhibition of expression is useful for
PT   treating and regulating bladder cancer-associated pathologies  -
XX
PS   Claim 4; Page 59; 64pp; English.
XX
CC   The sequence represents the coding sequence of bladder cancer-associated
CC   sequence, transitional cell carcinoma clone, TCC96C7. The sequence is
CC   upregulated in bladder cancer and its expression is indicative of bladder
CC   cancer. The sequence can be used as a marker, and can be used for
CC   diagnosing bladder cancer. Antibodies, ribozymes, antisense
CC   oligonucleotides or a dominant negative peptide directed against the
CC   sequence are useful for regulating bladder cancer-associated pathologies
CC   in a patient.
XX
SQ   Sequence 227 BP; 76 A; 63 C; 43 G; 45 T; 0 other;

  Query Match             27.3%;  Score 220;  DB 22;  Length 227;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-95;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     573 ATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACC 632
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGGCAAGCACCCAAGGGTGGCTGACC 60

Qy     633 AAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAG 692
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAG 120

Qy     693 CATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAG 752
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAG 180

Qy     753 CATCCAAACAAACAAGTAATAAATAAATAAATAAACTCAA 792
           |||||||||||||||||||||||||||||||||||||||
Db     181 CATCCAAACAAACAAGTAATAAATAAATAAATAAACTCAA 220


RESULT 11
ABN94440/c
ID   ABN94440 standard; DNA; 366 BP.
XX
AC   ABN94440;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #938 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US30589.
XX
PR   02-OCT-2000; 2000US-237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer,
PT   hepatocellular carcinoma or metastatic liver tumor in a patient,
PT   involves detecting the level of expression of two or more genes in a
PT   liver tissue sample  -
XX
PS   Claim 1; SEQ ID NO 938; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 366 BP; 77 A; 80 C; 119 G; 90 T; 0 other;

  Query Match             24.3%;  Score 196;  DB 24;  Length 366;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-83;
  Matches  246;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     357 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 307

Qy     417 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 247

Qy     477 TGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 536
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     246 TGAGCCAACCTCAACCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 187

Qy     537 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 596
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 127

Qy     597 GAGTGCA 603
           |||||||
Db     126 GAGTGCA 120


RESULT 12
ABL62262/c
ID   ABL62262 standard; DNA; 366 BP.
XX
AC   ABL62262;
```

XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:599.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
PS Claim 1; SEQ ID 599; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 366 BP; 77 A; 80 C; 119 G; 90 T; 0 other;

Query Match 24.3%; Score 196; DB 24; Length 366;
Best Local Similarity 99.6%; Pred. No. 1.5e-83;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    357 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 307

Qy    417 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 247

Qy    477 TGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 536
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    246 TGAGCCAACCTCAACCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 187

Qy    537 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 596
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 127

Qy    597 GAGTGCA 603
          |||||||
Db    126 GAGTGCA 120
```

RESULT 13
ABL67752/c
ID ABL67752 standard; DNA; 366 BP.
XX
AC ABL67752;
XX
DT 15-MAY-2002 (first entry)
XX
DE Oesophagus cancer related gene sequence SEQ ID NO:6089.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, 'Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 6089; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 366 BP; 77 A; 80 C; 119 G; 90 T; 0 other;

Query Match 24.3%; Score 196; DB 24; Length 366;
Best Local Similarity 99.6%; Pred. No. 1.5e-83;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    357 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTG 307

Qy    417 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 GGAAAGTTCCAGAACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCA 247

Qy    477 TGAGCCAACCTCACCCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 536
          |||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 TGAGCCAACCTCAACCCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAAC 187

Qy    537 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 596
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 CACCGAGCATTCCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTA 127

Qy    597 GAGTGCA 603
          |||||||
Db    126 GAGTGCA 120
```

RESULT 14
ABQ57070/c
ID ABQ57070 standard; cDNA; 629 BP.
XX
AC ABQ57070;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:765.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30732.
XX
PR 02-OCT-2000; 2000US-237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiaglingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy -
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the

```
CC   corresponding gene resides, and in tissue profiling, forensics, genetic
CC   analysis, mapping and diagnostic applications. (I) can be used to raise
CC   antibodies, and to screen for peptide analogues and antagonists.
XX
SQ   Sequence 629 BP; 128 A; 113 C; 149 G; 198 T; 41 other;

  Query Match             16.5%;  Score 133;  DB 24;  Length 629;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-53;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     633 AAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAG 692
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167 AAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTCAGCCTCCTGGCATTTAACTACCAG 108

Qy     693 CATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAG 752
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     107 CATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGCTGGAAAGAG 48

Qy     753 CATCCAAACAAAC 765
           |||||||||||||
Db      47 CATCCAAACAAAC 35


RESULT 15
AAD27026
ID   AAD27026 standard; DNA; 4159 BP.
XX
AC   AAD27026;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human Secreted epithelial colon stromal-1 (Secs-1) gene.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4159
FT                   /*tag=  a
FT                   /product= "Human Secs-1 protein"
FT                   /note= "CDS does not include stop codon"
FT                   /partial
FT   exon            1..69
FT                   /*tag=  d
FT                   /note= "Encodes Secs-1 peptide (AAE16484)"
FT   intron          70..2626
FT                   /*tag=  d
FT                   /cons_splice= (5'site:NO, 3'site:NO)
FT   exon            2627..2725
FT                   /*tag=  d
FT                   /note= "Encodes Secs-1 peptide (AAE16485)"
FT   intron          2726..4075
FT                   /*tag=  d
FT                   /cons_splice= (5'site:NO, 3'site:NO)
FT   exon            4076..4159
FT                   /*tag=  d
FT                   /note= "Encodes Secs-1 peptide (AAE16486)"
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
DR   P-PSDB; AAE16481, AAE16484, AAE16485, AAE16486.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Disclosure; Fig 4; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is human Secs-1
CC   gene.
XX
SQ   Sequence 4159 BP; 1009 A; 1037 C; 983 G; 1110 T; 20 other;

  Query Match             11.5%;  Score 93;  DB 24;  Length 4159;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-34;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      97 AGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCG 156
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2634 AGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCG 2693

Qy     157 AGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 189
           |||||||||||||||||||||||||||||||||
Db    2694 AGTCCCTAGCCCCAACTCAACAAACCTGAAAGG 2726


Search completed: May 17, 2003, 15:58:41
Job time : 276 secs
```

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2003, 03:39:29 ; Search time 1728 Seconds
(without alignments)
759.163 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09724000/runat_16052003_145500_17160/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09724000_@CGN_1_1_1906_@runat_16052003_145500_17160 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	442	100.0	356	10	AW351839	AW351839 RC1-CT019
2	418	94.6	503	9	AA422178	AA422178 zv31g07.r
3	416	94.1	906	12	BE899580	BE899580 601682443
4	356	80.5	557	14	BQ189412	BQ189412 UI-E-EJ1-
5	350.5	79.3	399	9	AA283751	AA283751 zt19g05.r
6	245	55.4	506	13	BM446173	BM446173 1IL4D7.ab
7	245	55.4	537	12	BF041606	BF041606 BP250007A
8	228.5	51.7	342	9	AA265120	AA265120 mz46g10.r
9	228.5	51.7	436	9	AA840147	AA840147 ud01f07.r
10	228.5	51.7	441	9	AA272330	AA272330 vb62a03.r
11	228.5	51.7	442	9	AA242158	AA242158 my30d03.r
12	228.5	51.7	447	9	AA265314	AA265314 mz47f10.r
13	228.5	51.7	450	10	BB864020	BB864020 BB864020
14	228.5	51.7	451	10	AW989615	AW989615 uf18c02.y
15	228.5	51.7	452	9	AA222093	AA222093 my29b11.r
16	228.5	51.7	456	9	AA267641	AA267641 mz52e07.r
17	228.5	51.7	456	9	AA529197	AA529197 vi33e01.r
18	228.5	51.7	462	9	AA017989	AA017989 mh48h07.r
19	228.5	51.7	498	9	AA240277	AA240277 my24c03.r
20	228.5	51.7	515	9	AA265055	AA265055 mz48g10.r
21	228.5	51.7	515	9	AA274151	AA274151 vb71f01.r
22	228.5	51.7	517	9	AA230541	AA230541 my30h12.r
23	228.5	51.7	521	9	AA272697	AA272697 vb62c09.r
24	228.5	51.7	523	9	AA272358	AA272358 vb62e04.r
25	228.5	51.7	543	9	AA538434	AA538434 vi99e06.r
26	228.5	51.7	549	9	AA220337	AA220337 my27g07.r
27	228.5	51.7	554	9	AA538362	AA538362 vi99a01.r
28	228.5	51.7	563	9	AA238890	AA238890 my36d01.r
29	228.5	51.7	753	13	BG966710	BG966710 602834571
30	228.5	51.7	758	11	AK012157	AK012157 Mus muscu
31	228.5	51.7	800	14	BQ939171	BQ939171 AGENCOURT
32	225.5	51.0	456	9	AA498458	AA498458 vh40c02.r
33	225.5	51.0	907	12	BF531305	BF531305 602091449
34	220.5	49.9	742	13	BG968801	BG968801 602836247
35	220.5	49.9	765	12	BF580962	BF580962 602100664
36	212.5	48.1	444	9	AA469817	AA469817 vf71a05.r
37	205.5	46.5	267	9	AA509491	AA509491 vg19c11.r
38	205.5	46.5	304	14	W08316	W08316 mb41d05.r1
39	199.5	45.1	529	9	AA463145	AA463145 vg85e06.r
40	199	45.0	422	10	AW142726	AW142726 EST292978
41	195	44.1	492	10	AW913863	AW913863 EST292204
42	195	44.1	510	14	BQ703407	BQ703407 OKST ERA0
43	195	44.1	531	12	BG374322	BG374322 UI-R-CV1-
44	194.5	44.0	466	9	AA528993	AA528993 vi30f09.r
45	182.5	41.3	1044	12	BF582268	BF582268 602094287

ALIGNMENTS

RESULT 1
AW351839
LOCUS AW351839 356 bp mRNA linear EST 01-FEB-2000
DEFINITION RC1-CT0199-150999-021-A03 CT0199 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW351839
VERSION AW351839.1 GI:6849552
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 356)
AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)

```
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC1&t2=RC1-CT0199-
            150999-021-A03&t3=1999-09-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 356.
FEATURES             Location/Qualifiers
     source          1. .356
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="CT0199"
                     /dev_stage="Adult"
                     /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT       80 a    117 c     86 g     73 t
ORIGIN

Alignment Scores:
Pred. No.:              8.53e-37        Length:         356
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     10              Gaps:           0

US-09-724-000A-5 (1-81) x AW351839 (1-356)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 114

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 174

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     175 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 234

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     235 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 294

Qy      81 Val 81
           |||
Db     295 GTG 297

RESULT 2
AA422178
LOCUS       AA422178                503 bp    mRNA    linear   EST 16-OCT-1997
DEFINITION  zv31g07.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:755292 5', mRNA sequence.
ACCESSION   AA422178
VERSION     AA422178.1  GI:2101029
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 503.
FEATURES             Location/Qualifiers
     source          1. .503
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:755292"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT      108 a    167 c    119 g    109 t
ORIGIN

Alignment Scores:
Pred. No.:              4.68e-34        Length:         503
Score:                  418.00          Matches:        79
Percent Similarity:     98.77%          Conservative:   1
Best Local Similarity:  97.53%          Mismatches:     1
Query Match:            94.57%          Indels:         1
DB:                     9               Gaps:           0

US-09-724-000A-5 (1-81) x AA422178 (1-503)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 75

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76 TCCACAGAGGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 135

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     136 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 195

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||   :::||||||
Db     196 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGA-GTCCCACAG 254

Qy      81 Val 81
           |||
Db     255 GTG 257

RESULT 3
BE899580
LOCUS       BE899580                906 bp    mRNA    linear   EST 29-SEP-2000
DEFINITION  601682443F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3952523 5',
            mRNA sequence.
ACCESSION   BE899580
```

```
VERSION     BE899580.1  GI:10367234
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM822  row: f  column: 12
            High quality sequence start: 23
            High quality sequence stop: 775.
FEATURES             Location/Qualifiers
     source          1. .906
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3952523"
                     /clone_lib="NIH_MGC_9"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      219 a    282 c    219 g    186 t
ORIGIN

Alignment Scores:
Pred. No.:              1.78e-33        Length:         906
Score:                  416.00          Matches:        81
Percent Similarity:     97.59%          Conservative:   0
Best Local Similarity:  97.59%          Mismatches:     0
Query Match:            94.12%          Indels:         2
DB:                     12              Gaps:           0

US-09-724-000A-5 (1-81) x BE899580 (1-906)

Qy      1 MetArgLeuLeu-ValLeuSerSerLeuLeu-CysIleLeuLeuLeuCysPheSerIleP 20
          ||||||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db     69 ATGAGGCTTCTAGGTCCTTTCCAGCCTGCTTCTGTATCCTGCTTCTCTGCTTCTCCATCT 128

Qy     20 heSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuC 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 TCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCT 188

Qy     40 ysCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 GCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCT 248

Qy     60 ysLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProG 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 GTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCAC 308

Qy     80 lnVal 81
          |||||
Db    309 AGGTG 313

RESULT 4
BQ189412
LOCUS       BQ189412                 557 bp    mRNA    linear   EST 30-APR-2002
DEFINITION  UI-E-EJ1-ajv-h-16-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
            UI-E-EJ1-ajv-h-16-0-UI 5', mRNA sequence.
ACCESSION   BQ189412
VERSION     BQ189412.1  GI:20364963
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 557)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 REVERSE.
FEATURES             Location/Qualifiers
     source          1. .557
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="UI-E-EJ1-ajv-h-16-0-UI"
                     /clone_lib="UI-E-EJ1"
                     /tissue_type="fetal eyes, lens, eye anterior segment,
                     optic nerve, retina, Retina Foveal and Macular, RPE and
                     Choroid"
                     /dev_stage="fetal and adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-EJ1 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are: fetal eyes, AGAATCAAGA
                     ; lens, CGATTAGCGA; eye anterior segment, AATGCCGCAT;
                     optic nerve, CCATTAAGTG; retina, CCGCG; Retina Foveal and
                     Macular, GTCC; RPE and Choroid, ACCTA. This library was
                     created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."
BASE COUNT      135 a    182 c    138 g    101 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.78e-27        Length:         557
Score:                  356.00          Matches:        63
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            80.54%          Indels:         0
DB:                     14              Gaps:           0

US-09-724-000A-5 (1-81) x BQ189412 (1-557)

Qy     19 IlePheSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArg 38
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATCTTCTCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGG 60
```

```
Qy      39 LeuCysCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGG 120

Qy      59 LeuCysLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeu 78
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTC 180

Qy      79 ProGlnVal 81
           |||||||||
Db     181 CCACAGGTG 189

RESULT 5
AA283751
LOCUS       AA283751                399 bp    mRNA    linear   EST 08-AUG-1997
DEFINITION  zt19g05.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:713624 5', mRNA sequence.
ACCESSION   AA283751
VERSION     AA283751.1  GI:1928032
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 857   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 347.
FEATURES             Location/Qualifiers
     source          1. .399
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:713624"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT       89 a    127 c     97 g     86 t
ORIGIN

Alignment Scores:
Pred. No.:              4.16e-27        Length:         399
Score:                  350.50          Matches:        72
Percent Similarity:     90.24%          Conservative:   2
Best Local Similarity:  87.80%          Mismatches:     5
Query Match:            79.30%          Indels:         4
DB:                     9               Gaps:           1

US-09-724-000A-5 (1-81) x AA283751 (1-399)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 92

Qy      21 SerThrGluGlyLysArgArgProAlaLys-AlaTrpSerGlyArgArgThrArgLeuCy 40
           ||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db      93 TCCACAGAAGGGAAGAGGCGTCCTGCCAACAGCCTGGTCAGGCAGGAGAACCAGGCTCTG 152

Qy      40 sCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCy 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 CTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTG 212

Qy      60 sLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGl 80
           ||||||||||||||||   |||||||||   |||   |||:::          :::||||||
Db     213 TAAACCATGCAAGCTTGA-CCAGAGCCCGCTTTGGCTGTGCTG------GGGATCCCACA 265

Qy      80 nVal 81
           ||||
Db     266 GGTG 269

RESULT 6
BM446173
LOCUS       BM446173                506 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  1IL4D7.ab1 Bos taurus Ileum #1 library Bos taurus cDNA, mRNA
            sequence.
ACCESSION   BM446173
VERSION     BM446173.1  GI:18530329
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 506)
  AUTHORS   Hansen,C., Fu,A., Meng,Y., Li,C., Okine,E., Sensen,C.W., Gordon
            ,P.M.K. and Moore,S.S.
  TITLE     Gene Expression Profiling of the Bovine Gastrointestinal Tract
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Dr. Stephen Moore
            . Beef Genomics Laboratory
            Dept of AFNS, University of Alberta
            410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
            Tel: 780 492 0169
            Fax: 780 492 4265
            Email: smoore@afns.ualberta.ca
            Insert Length: 506   Std Error: 0.00
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .506
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="Bos taurus Ileum #1 library"
                     /tissue_type="Smooth muscle"
                     /cell_type="Simple columnar epithelial"
                     /dev_stage="Young adult"
                     /lab_host="XL1-BlueMRF'strain"
                     /note="Organ: Intestine/ileum; Vector: Uni-2ZAPXR; Site_1:
                     EcoRI; Site_2: Xho I"
BASE COUNT      115 a    162 c    113 g    116 t
ORIGIN

Alignment Scores:
Pred. No.:              7.17e-16        Length:         506
Score:                  245.00          Matches:        48
Percent Similarity:     69.05%          Conservative:   10
Best Local Similarity:  57.14%          Mismatches:     20
Query Match:            55.43%          Indels:         6
DB:                     13              Gaps:           2
```

```
US-09-724-000A-5 (1-81) x BM446173 (1-506)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db      39 ATGAGGCTTCTAATCCTCACCAGCCTGCTCTGCATCCTGCTTCTCTGCTTGTCCGTCTTC 98

Qy      21 SerThrGluGlyLys---------ArgArgProAlaLysAlaTrpSerGlyArgArgThr 37
Db      99 TCCGCAGAAGGGAGAAGTCACCTGAGGCATCATGCCAAGCCCGGAAAAGGCAAGCCC--- 155

Qy      38 ArgLeuCysCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisVal 57
Db     156 ------TGCTGTCCCCGAATTCCTGGCCCTGACCTGATGACCCAGAAAGGACACCGTACG 209

Qy      58 ArgLeuCysLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAla 77
Db     210 AGAAACTGCAGACCATGCAAACTCAAGTCAAAGCACCGCTTTTGGGTGGTTCCTGGGGCA 269

Qy      78 LeuProGlnVal 81
Db     270 CTCCCACAGGTG 281

RESULT 7
BF041606
LOCUS       BF041606                537 bp    mRNA    linear   EST 10-OCT-2000
DEFINITION  BP250007A20H3 Soares normalized bovine placenta Bos taurus cDNA
            clone BP250007A20H3 5', mRNA sequence.
ACCESSION   BF041606
VERSION     BF041606.1  GI:10758652
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson
            ,J.H.
  TITLE     Bovine ESTs
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Lewin, H. A.
            W. M. Keck Center for Comparative and Functional Genomics
            University of Illinois at Urbana-Champaign
            340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
            61801, USA
            Tel: 217 333 5998
            Fax: 217 244 5617
            Email: h-lewin@uiuc.edu
            Funding for cattle EST sequencing was provided by the USDA National
            Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
            to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
            from Washington University Genome Center. Vector Trimmi g:
            Cross_match from Washington University Genome Center PHRAP suite.
            This sequence is vector free and at least 200 bp in length.
            PCR PRimers
            FORWARD: TAATACGACTCACTATAGGG
            BACKWARD: ATTAACCCTCACTAAAG
            Insert Length: 537   Std Error: 0.00
            Plate: BP250007A20  row: H  column: 3
            Seq primer: AGCGGATAACAATTTCACACAGGA
            High quality sequence stop: 537.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="BP250007A20H3"
                     /clone_lib="Soares normalized bovine placenta"
                     /sex="female"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3Pac; Site_1: EcoRI;
                     Site_2: NotI; The cDNA library was contributed by the
                     Soares laboratory and it  was constructed and normalized
                     as described by Bonaldo, M.F.,  Lennon, G. and Soares,
                     M.B. (1996), Genome Research 6(9): 791-806. "
BASE COUNT      128 a    166 c    120 g    122 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              7.81e-16        Length:         537
Score:                  245.00          Matches:        48
Percent Similarity:     69.05%          Conservative:   10
Best Local Similarity:  57.14%          Mismatches:     20
Query Match:            55.43%          Indels:         6
DB:                     12              Gaps:           2

US-09-724-000A-5 (1-81) x BF041606 (1-537)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db      28 ATGAGGCTTCTAATCCTCACCAGCCTGCTCTGCATCCTGCTTCTCTGCTTGTCCGTCTTC 87

Qy      21 SerThrGluGlyLys---------ArgArgProAlaLysAlaTrpSerGlyArgArgThr 37
Db      88 TCCGCAGAAGGGAGAAGTCACCTGAGGCATCATGCCAAGCCCGGAAAAGGCAAGCCC--- 144

Qy      38 ArgLeuCysCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisVal 57
Db     145 ------TGCTGTCCCCGAATTCCTGGCCCTGACCTGATGACCCAGAAAGGACACCGTACG 198

Qy      58 ArgLeuCysLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAla 77
Db     199 AGAAACTGCAGACCATGCAAACTCAAGTCAAAGCACCGCTTTTGGGTGGTTCCTGGGGCA 258

Qy      78 LeuProGlnVal 81
Db     259 CTCCCACAGGTG 270

RESULT 8
AA265120
LOCUS       AA265120                342 bp    mRNA    linear   EST 20-MAR-1997
DEFINITION  mz46g10.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
            clone IMAGE:716514 5', mRNA sequence.
ACCESSION   AA265120
VERSION     AA265120.1  GI:1901208
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 342)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:442010
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 329.
FEATURES             Location/Qualifiers
     source          1. .342
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:716514"
                     /clone_lib="Barstead mouse pooled organs MPLRB4"
```

```
                     /sex="mixed"
                     /tissue_type="pooled organs"
                     /dev_stage="7 day"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [GTTGGATTCGGTACC], digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
BASE COUNT       83 a    103 c     78 g     78 t
ORIGIN

Alignment Scores:
Pred. No.:              2.21e-14        Length:         342
Score:                  228.50          Matches:        45
Percent Similarity:     67.90%          Conservative:   10
Best Local Similarity:  55.56%          Mismatches:     23
Query Match:            51.70%          Indels:         3
DB:                     9               Gaps:           1

US-09-724-000A-5 (1-81) x AA265120 (1-342)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db     37 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 96

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          |||:::||||||:::|||   |||||||||:::         ||||||         |||
Db     97 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 147

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||      |||             |||      ||||||:::|||   |||   |||
Db    148 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 207

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          :::   |||:::   :::         :::   ||||||||||||||||||||||||||||
Db    208 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 267

Qy     81 Val 81
          :::
Db    268 ATA 270

RESULT 9
AA840147
LOCUS       AA840147                436 bp    mRNA    linear   EST 27-FEB-1998
DEFINITION  ud01f07.r1 Soares_NMPu Mus musculus cDNA clone IMAGE:1433893 5',
            mRNA sequence.
ACCESSION   AA840147
VERSION     AA840147.1  GI:2915856
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 436)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:917961
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 422.
FEATURES             Location/Qualifiers
     source          1. .436
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1433893"
                     /clone_lib="Soares_NMPu"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; 1st strand cDNA was prepared from
                     pregnant mouse uterus, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      105 a    130 c     99 g    102 t
ORIGIN

Alignment Scores:
Pred. No.:              3.13e-14        Length:         436
Score:                  228.50          Matches:        45
Percent Similarity:     67.90%          Conservative:   10
Best Local Similarity:  55.56%          Mismatches:     23
Query Match:            51.70%          Indels:         3
DB:                     9               Gaps:           1

US-09-724-000A-5 (1-81) x AA840147 (1-436)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db     13 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 72

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          |||:::||||||:::|||   |||||||||:::         ||||||         |||
Db     73 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 123

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||      |||             |||      ||||||:::|||   |||   |||
Db    124 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 183

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          :::   |||:::   :::         :::   ||||||||||||||||||||||||||||
Db    184 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 243

Qy     81 Val 81
          :::
Db    244 ATA 246

RESULT 10
AA272330
LOCUS       AA272330                441 bp    mRNA    linear   EST 26-MAR-1997
DEFINITION  vb62a03.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
            clone IMAGE:761548 5', mRNA sequence.
ACCESSION   AA272330
VERSION     AA272330.1  GI:1910661
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 441)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
```

```
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:462468
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 404.
FEATURES             Location/Qualifiers
     source          1. .441
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:761548"
                     /clone_lib="Barstead mouse pooled organs MPLRB4"
                     /sex="mixed"
                     /tissue_type="pooled organs"
                     /dev_stage="7 day"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [GTTGGATTCGGTACC], digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
BASE COUNT      103 a    133 c    101 g    104 t
ORIGIN

Alignment Scores:
Pred. No.:              3.19e-14        Length:         441
Score:                  228.50          Matches:        45
Percent Similarity:     67.90%          Conservative:   10
Best Local Similarity:  55.56%          Mismatches:     23
Query Match:            51.70%          Indels:         3
DB:                     9               Gaps:           1

US-09-724-000A-5 (1-81) x AA272330 (1-441)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db      37 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 96

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           |||:::||||||:::|||   |||||||||:::         ||||||         |||
Db      97 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 147

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||      |||          |||      ||||||:::|||   |||   |||
Db     148 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 207

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           :::   |||:::   :::          :::   ||||||||||||||||||||||||||||
Db     208 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 267

Qy      81 Val 81
           :::
Db     268 ATA 270

RESULT 11
AA242158
LOCUS       AA242158                442 bp    mRNA    linear   EST 07-MAR-1997
DEFINITION  my30d03.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
            clone IMAGE:697349 5', mRNA sequence.
ACCESSION   AA242158
VERSION     AA242158.1  GI:1873383
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:430909
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 388.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:697349"
                     /clone_lib="Barstead mouse pooled organs MPLRB4"
                     /sex="mixed"
                     /tissue_type="pooled organs"
                     /dev_stage="7 day"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [GTTGGATTCGGTACC], digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
BASE COUNT      103 a    134 c    102 g    103 t
ORIGIN

Alignment Scores:
Pred. No.:              3.2e-14         Length:         442
Score:                  228.50          Matches:        45
Percent Similarity:     67.90%          Conservative:   10
Best Local Similarity:  55.56%          Mismatches:     23
Query Match:            51.70%          Indels:         3
DB:                     9               Gaps:           1

US-09-724-000A-5 (1-81) x AA242158 (1-442)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db      52 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 111

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           |||:::||||||:::|||   |||||||||:::         ||||||         |||
Db     112 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 162

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||      |||          |||      ||||||:::|||   |||   |||
Db     163 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 222

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           :::   |||:::   :::          :::   ||||||||||||||||||||||||||||
Db     223 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 282

Qy      81 Val 81
```

```
          :::
Db    283 ATA 285

RESULT 12
AA265314
LOCUS       AA265314                447 bp    mRNA    linear   EST 20-MAR-1997
DEFINITION  mz47f10.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
            clone IMAGE:716587 5', mRNA sequence.
ACCESSION   AA265314
VERSION     AA265314.1  GI:1901400
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:442083
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 393.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:716587"
                     /clone_lib="Barstead mouse pooled organs MPLRB4"
                     /sex="mixed"
                     /tissue_type="pooled organs"
                     /dev_stage="7 day"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [GTTGGATTCGGTACC], digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
BASE COUNT      105 a    136 c    101 g    105 t
ORIGIN

Alignment Scores:
Pred. No.:              3.25e-14        Length:         447
Score:                  228.50          Matches:        45
Percent Similarity:     67.90%          Conservative:   10
Best Local Similarity:  55.56%          Mismatches:     23
Query Match:            51.70%          Indels:         3
DB:                     9               Gaps:           1

US-09-724-000A-5 (1-81) x AA265314 (1-447)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db     49 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 108

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          |||:::||||||:::|||   ||||||||||:::         ||||||          |||
          T
```

```
Db    109 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 159

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||      |||          |||      ||||||:::|||   |||   |||
Db    160 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 219

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          :::   |||:::   :::         :::   ||||||||||||||||||||||||||||
Db    220 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 279

Qy     81 Val 81
          :::
Db    280 ATA 282

RESULT 13
BB864020
LOCUS       BB864020                450 bp    mRNA    linear   EST 27-NOV-2001
DEFINITION  BB864020 RIKEN full-length enriched, kidney CCL-142 RAG cDNA Mus
            musculus cDNA clone G430041F19 5', mRNA sequence.
ACCESSION   BB864020
VERSION     BB864020.1  GI:17110230
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 450)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .450
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="G430041F19"
                     /clone_lib="RIKEN full-length enriched, kidney CCL-142 RAG
```

cDNA"
/tissue_type="kidney"
/cell_line="CCL-142 RAG"
/note="pooled cell lines ; (cell_line=CRL-1751 WEHI 164), (cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3), (cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA), (cell_line=RCB-0559 K-1 . F1), (cell_line=RCB-1283 B16 melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231), (cell_type=Leydig cells, cell_line=CRL-2065 MLTC-1), (cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE), (tissue_type=bladder, cell_line=RCB-0544 MBT-2), (tissue_type=bone marrow, cell_type=stroma cell, cell_line= CRL-2028 SR-4987), (tissue_type=colon, cell_line=RCB-0549 Cle-H3), (tissue_type=kidney, cell_line=CCL-142 RAG), (tissue_type=submandibular gland, cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C, cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3), (strain=C3H, tissue_type=brain, cell_line=CRL-1443 BC3H1)"

BASE COUNT 101 a 136 c 107 g 106 t
ORIGIN

Alignment Scores:

Pred. No.:	3.28e-14	Length:	450
Score:	228.50	Matches:	45
Percent Similarity:	67.90%	Conservative:	10
Best Local Similarity:	55.56%	Mismatches:	23
Query Match:	51.70%	Indels:	3
DB:	10	Gaps:	1

US-09-724-000A-5 (1-81) x BB864020 (1-450)

```
Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db      57 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 116

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           |||:::||||||:::|||   |||||||||:::         ||||||         |||
Db     117 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 167

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||      |||             |||      ||||||:::|||   |||   |||
Db     168 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 227

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           :::   |||:::   :::         :::   ||||||||||||||||||||||||||||
Db     228 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 287

Qy      81 Val 81
           :::
Db     288 ATA 290
```

RESULT 14
AW989615
LOCUS AW989615 451 bp mRNA linear EST 02-JUN-2000
DEFINITION uf18c02.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone IMAGE:1511714 5', mRNA sequence.
ACCESSION AW989615
VERSION AW989615.1 GI:8185064
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 451)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:938566
Seq primer: -40RP from Gibco
High quality sequence stop: 451.

FEATURES Location/Qualifiers
source 1. .451
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1511714"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 105 a 137 c 102 g 107 t
ORIGIN

Alignment Scores:

Pred. No.:	3.29e-14	Length:	451
Score:	228.50	Matches:	45
Percent Similarity:	67.90%	Conservative:	10
Best Local Similarity:	55.56%	Mismatches:	23
Query Match:	51.70%	Indels:	3
DB:	10	Gaps:	1

US-09-724-000A-5 (1-81) x AW989615 (1-451)

```
Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db      38 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 97

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           |||:::||||||:::|||   |||||||||:::         ||||||         |||
Db      98 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 148

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           ||||||      |||             |||      ||||||:::|||   |||   |||
Db     149 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 208

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           :::   |||:::   :::         :::   ||||||||||||||||||||||||||||
Db     209 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 268

Qy      81 Val 81
           :::
Db     269 ATA 271
```

RESULT 15
AA222093
LOCUS AA222093 452 bp mRNA linear EST 14-FEB-1997
DEFINITION my29b11.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA clone IMAGE:697245 5', mRNA sequence.
ACCESSION AA222093
VERSION AA222093.1 GI:1841618
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 452)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)

```
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:430805
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 420.
FEATURES             Location/Qualifiers
     source          1. .452
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:697245"
                     /clone_lib="Barstead mouse pooled organs MPLRB4"
                     /sex="mixed"
                     /tissue_type="pooled organs"
                     /dev_stage="7 day"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [GTTGGATTCGGTACC], digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
BASE COUNT      107 a    135 c    104 g    106 t
ORIGIN

Alignment Scores:
Pred. No.:              3.3e-14         Length:         452
Score:                  228.50          Matches:        45
Percent Similarity:     67.90%          Conservative:   10
Best Local Similarity:  55.56%          Mismatches:     23
Query Match:            51.70%          Indels:         3
DB:                     9               Gaps:           1

US-09-724-000A-5 (1-81) x AA222093 (1-452)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db     49 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 108

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          |||:::||||||:::|||   |||||||||:::         ||||||         |||
Db    109 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 159

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||      |||             |||      ||||||:::|||   |||   |||
Db    160 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 219

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          :::   |||:::   :::         :::   ||||||||||||||||||||||||||
Db    220 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 279

Qy     81 Val 81
          :::
Db    280 ATA 282


Search completed: May 17, 2003, 12:22:42
Job time : 1736 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:35:10 ; Search time 20 Seconds
(without alignments)
390.626 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 362588 seqs, 96450795 residues

Word size : 0

Total number of hits satisfying chosen parameters: 47966

Minimum DB seq length: 23
Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6	7.4	28	9	US-09-749-637A-265	Sequence 265, App
2	6	7.4	54	10	US-09-864-761-47036	Sequence 47036, A
3	6	7.4	65	9	US-10-091-504-786	Sequence 786, App
4	6	7.4	65	10	US-09-764-869-786	Sequence 786, App
5	6	7.4	72	10	US-09-864-761-36193	Sequence 36193, A
6	6	7.4	78	10	US-09-864-761-37299	Sequence 37299, A
7	6	7.4	79	9	US-10-083-357-1164	Sequence 1164, Ap
8	5	6.2	23	8	US-08-424-550B-400	Sequence 400, App
9	5	6.2	24	9	US-09-843-221A-132	Sequence 132, App
10	5	6.2	24	9	US-09-843-221A-142	Sequence 142, App
11	5	6.2	24	9	US-10-150-111-127	Sequence 127, App
12	5	6.2	24	10	US-09-242-980-15	Sequence 15, Appl
13	5	6.2	24	12	US-10-001-879-157	Sequence 157, App
14	5	6.2	25	9	US-09-989-920-214	Sequence 214, App
15	5	6.2	26	10	US-09-925-300-1735	Sequence 1735, Ap
16	5	6.2	27	9	US-10-083-357-1099	Sequence 1099, Ap
17	5	6.2	27	10	US-09-864-761-39998	Sequence 39998, A
18	5	6.2	27	10	US-09-864-761-40419	Sequence 40419, A
19	5	6.2	27	10	US-09-911-888-8	Sequence 8, Appli
20	5	6.2	28	9	US-09-843-221A-96	Sequence 96, Appl
21	5	6.2	28	9	US-09-843-221A-106	Sequence 106, App
22	5	6.2	28	10	US-09-864-761-41850	Sequence 41850, A
23	5	6.2	29	9	US-09-362-179-5	Sequence 5, Appli
24	5	6.2	29	10	US-09-864-761-44425	Sequence 44425, A
25	5	6.2	30	9	US-09-843-221A-127	Sequence 127, App
26	5	6.2	30	9	US-09-843-221A-137	Sequence 137, App
27	5	6.2	30	9	US-09-843-221A-156	Sequence 156, App
28	5	6.2	30	12	US-10-127-318-5	Sequence 5, Appli
29	5	6.2	31	10	US-09-142-755-10	Sequence 10, Appl
30	5	6.2	32	9	US-09-983-802-207	Sequence 207, App
31	5	6.2	32	10	US-09-922-261-157	Sequence 157, App
32	5	6.2	32	10	US-09-049-695A-20	Sequence 20, Appl
33	5	6.2	33	10	US-09-975-901-17	Sequence 17, Appl
34	5	6.2	34	9	US-09-843-221A-91	Sequence 91, Appl
35	5	6.2	34	9	US-09-843-221A-101	Sequence 101, App
36	5	6.2	34	9	US-09-843-221A-120	Sequence 120, App
37	5	6.2	35	10	US-09-864-761-44044	Sequence 44044, A
38	5	6.2	35	10	US-09-922-261-155	Sequence 155, App
39	5	6.2	36	9	US-09-764-891-2801	Sequence 2801, Ap
40	5	6.2	36	10	US-09-864-761-35960	Sequence 35960, A
41	5	6.2	36	10	US-09-864-761-36141	Sequence 36141, A
42	5	6.2	36	10	US-09-867-852-20	Sequence 20, Appl
43	5	6.2	37	10	US-09-242-980-14	Sequence 14, Appl
44	5	6.2	38	9	US-09-984-245-207	Sequence 207, App
45	5	6.2	38	9	US-09-966-262-207	Sequence 207, App

ALIGNMENTS

RESULT 1
US-09-749-637A-265
; Sequence 265, Application US/09749637A
; Patent No. US20020173449A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cartier, G. Edward
; APPLICANT: Watkins, Maren
; APPLICANT: Hillyard, David R.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Layer, Richard T.
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: O-Superfamily Conotoxin Peptides
; FILE REFERENCE: 2314-227
; CURRENT APPLICATION NUMBER: US/09/749,637A
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/243,412
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US60/219,440
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/214,263
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 60/173,754
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 409
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 265
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Conus delessertii
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(28)
; OTHER INFORMATION: Xaa at residues 4 and 14 may be Pro or hydroxy-Pro; Xaa at resid
; OTHER INFORMATION: e 7 may be Glu or gamma-carboxy-Glu; Xaa at residues 15 and 21 n
; OTHER INFORMATION: y be Tyr, 125-I-Tyr, mono-iodo-Tyr, di-iodo-Tyr, O-sulpho-Tyr or
; OTHER INFORMATION: O-phospho-Tyr
US-09-749-637A-265

Query Match 7.4%; Score 6; DB 9; Length 28;

```
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 CILLLC 16
           ||||||
Db      22 CILLLC 27


RESULT 2
US-09-864-761-47036
; Sequence 47036, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47036
;   LENGTH: 54
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005317.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.64
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.59
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.68
;   OTHER INFORMATION: EST_HUMAN HIT: AW500861.1, EVALUE 2.00e-15
US-09-864-761-47036

  Query Match             7.4%;  Score 6;  DB 10;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLVLSS 8
            ||||||
Db      35 LLVLSS 40


RESULT 3
US-10-091-504-786
; Sequence 786, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC007C1
;  CURRENT APPLICATION NUMBER: US/10/091,504
;  CURRENT FILING DATE:  2002-03-07
;  NUMBER OF SEQ ID NOS: 2442
;  Prior Application removed - See File Wrapper or Palm
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 786
;   LENGTH: 65
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (17)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: misc_feature
;   LOCATION: (58)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-091-504-786

  Query Match             7.4%;  Score 6;  DB 9;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLVLSS 8
            ||||||
Db       6 LLVLSS 11


RESULT 4
US-09-764-869-786
; Sequence 786, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC007
;  CURRENT APPLICATION NUMBER: US/09/764,869
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 2442
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 786
;   LENGTH: 65
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (17)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (58)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-786

  Query Match             7.4%;  Score 6;  DB 10;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      3 LLVLSS 8
          ||||||
Db      6 LLVLSS 11


RESULT 5
US-09-864-761-36193
; Sequence 36193, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36193
;   LENGTH: 72
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC012127.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
;   OTHER INFORMATION: SWISSPROT HIT: O61608, EVALUE 7.60e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BE789526.1, EVALUE 8.10e-01
US-09-864-761-36193

  Query Match             7.4%;  Score 6;  DB 10;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 LCILLL 15
             ||||||
Db       1 LCILLL 6


RESULT 6
US-09-864-761-37299
; Sequence 37299, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37299
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC010906.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
```

```
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =   3
US-09-864-761-37299

  Query Match             7.4%;  Score 6;  DB 10;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 LLVLSS 8
          ||||||
Db     71 LLVLSS 76


RESULT 7
US-10-083-357-1164
; Sequence 1164, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
;  APPLICANT: Qiandong Zeng et al.
;  TITLE OF INVENTION: Systemic Discovery of New Genes
;  FILE REFERENCE: 032796-090
;  CURRENT APPLICATION NUMBER: US/10/083,357
;  CURRENT FILING DATE:  2002-02-27
;  NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 1164
;   LENGTH: 79
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-10-083-357-1164

  Query Match             7.4%;  Score 6;  DB 9;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 RLLVLS 7
          ||||||
Db     67 RLLVLS 72


RESULT 8
US-08-424-550B-400
; Sequence 400, Application US/08424550B
; Patent No. US20020119447A1
;  GENERAL INFORMATION:
;    APPLICANT:  JOHN N. SIMONS
;    APPLICANT:  TAMI J. PILOT-MATIAS
;    APPLICANT:  GEORGE J. DAWSON
;    APPLICANT:  GEORGE G. SCHLAUDER
;    APPLICANT:  SURESH M. DESAI
;    APPLICANT:  THOMAS P. LEARY
;    APPLICANT:  ANTHONY SCOTT MUERHOFF
;    APPLICANT:  JAMES C. ERKER
;    APPLICANT:  SHERI L. BUIJK
;    APPLICANT:  ISA K. MUSHAHWAR
;    TITLE OF INVENTION:  NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
;    TITLE OF INVENTION:  REAGENTS AND METHODS FOR THEIR USE
;    NUMBER OF SEQUENCES:  716
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ABBOTT LABORATORIES  D377/AP6D
;      STREET:  100 ABBOTT PARK ROAD
;      CITY:  ABBOTT PARK
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/424,550B
;      FILING DATE:
;      CLASSIFICATION:  435435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  POREMBSKI, PRISCILLA E.
;      REGISTRATION NUMBER:  33,207
;      REFERENCE/DOCKET NUMBER:  5527.PC.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708-937-6365
;      TELEFAX:  708-938-2623
;  INFORMATION FOR SEQ ID NO:  400:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-424-550B-400

  Query Match             6.2%;  Score 5;  DB 8;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     24 GKRRP 28
          |||||
Db      8 GKRRP 12


RESULT 9
US-09-843-221A-132
; Sequence 132, Application US/09843221A
; Publication No. US20030039654A1
; GENERAL INFORMATION:
;  APPLICANT: KOSTENUIK, PAUL
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  LACEY, DAVID LEE
;  TITLE OF INVENTION: MODULATORS OF RECEPTORS FOR PARATHYROID HORMONE AND PARATHYROID
;  TITLE OF INVENTION:  RELATED PROTEIN
;  FILE REFERENCE: A-665B
;  CURRENT APPLICATION NUMBER: US/09/843,221A
;  CURRENT FILING DATE:  2001-04-26
;  PRIOR APPLICATION NUMBER: 60/266,673
;  PRIOR FILING DATE: 2001-02-06
;  PRIOR APPLICATION NUMBER: 60/214,860
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/200,053
;  PRIOR FILING DATE: 2000-04-27
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 132
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: modified human PTH
US-09-843-221A-132

  Query Match             6.2%;  Score 5;  DB 9;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 LSSLL 10
          |||||
Db     18 LSSLL 22


RESULT 10
US-09-843-221A-142
; Sequence 142, Application US/09843221A
; Publication No. US20030039654A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: KOSTENUIK, PAUL
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  LACEY, DAVID LEE
;  TITLE OF INVENTION: MODULATORS OF RECEPTORS FOR PARATHYROID HORMONE AND PARATHYROID H
;  TITLE OF INVENTION:  RELATED PROTEIN
;  FILE REFERENCE: A-665B
;  CURRENT APPLICATION NUMBER: US/09/843,221A
;  CURRENT FILING DATE:  2001-04-26
;  PRIOR APPLICATION NUMBER: 60/266,673
;  PRIOR FILING DATE: 2001-02-06
;  PRIOR APPLICATION NUMBER: 60/214,860
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/200,053
;  PRIOR FILING DATE: 2000-04-27
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 142
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: modified human PTHrP
US-09-843-221A-142

  Query Match             6.2%;  Score 5;  DB 9;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 LSSLL 10
          |||||
Db     18 LSSLL 22


RESULT 11
US-10-150-111-127
; Sequence 127, Application US/10150111
; Publication No. US20030078386A1
; GENERAL INFORMATION:
;  APPLICANT: Rubin et al.
;  TITLE OF INVENTION: Secreted Protein HPEAD48
;  FILE REFERENCE: PZ018P1D1
;  CURRENT APPLICATION NUMBER: US/10/150,111
;  CURRENT FILING DATE:  2002-05-20
;  PRIOR APPLICATION NUMBER: 09/288,143
;  PRIOR FILING DATE: 1999-04-08
;  PRIOR APPLICATION NUMBER: PCT/US98/21142
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 60/061,463
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,529
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/071,498
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,527
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,536
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,532
;  PRIOR FILING DATE: 1997-10-09
;  NUMBER OF SEQ ID NOS: 219
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 127
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-111-127

  Query Match             6.2%;  Score 5;  DB 9;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 LVLSS 8
          |||||
Db      9 LVLSS 13


RESULT 12
US-09-242-980-15
; Sequence 15, Application US/09242980
; Patent No. US20010021073A1
; GENERAL INFORMATION:
;  APPLICANT:  Ogi, Kasuhiro
;  APPLICANT:  Onda, Haruo
;  TITLE OF INVENTION:  NOVEL PROTEIN, ITS PRODUCTION AND USE
;  FILE REFERENCE:  2417US0P
;  CURRENT APPLICATION NUMBER: US/09/242,980
;  CURRENT FILING DATE:  1999-02-28
;  PRIOR APPLICATION NUMBER:  PCT/JP97/03194
;  PRIOR FILING DATE:  1997-08-10
;  PRIOR APPLICATION NUMBER:  JP 8-240880
;  PRIOR FILING DATE:  1996-09-11
;  PRIOR APPLICATION NUMBER:  JP 8-318049
;  PRIOR FILING DATE:  1996-11-28
;  PRIOR APPLICATION NUMBER:  JP 9-135633
;  PRIOR FILING DATE:  1997-05-27
;  NUMBER OF SEQ ID NOS:  35
;  SOFTWARE:
; SEQ ID NO 15
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Human
US-09-242-980-15

  Query Match             6.2%;  Score 5;  DB 10;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     74 VPGAL 78
          |||||
Db     10 VPGAL 14


RESULT 13
US-10-001-879-157
; Sequence 157, Application US/10001879
; Patent No. US20020127237A1
; GENERAL INFORMATION:
;  APPLICANT: Salceda, Susana
;  APPLICANT:  Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Cafferkey, Robert
;  APPLICANT:  Ali, Shujath
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
;  FILE REFERENCE: DEX-0281
;  CURRENT APPLICATION NUMBER: US/10/001,879
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 60/252,188
;  PRIOR FILING DATE: 2000-11-21
;  NUMBER OF SEQ ID NOS: 201
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 157
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-001-879-157

  Query Match             6.2%;  Score 5;  DB 12;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 LVLSS 8
          |||||
Db     12 LVLSS 16
```

```
RESULT 14
US-09-989-920-214
; Sequence 214, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Chen, Sei-Yu
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;  FILE REFERENCE: DEX-0291
;  CURRENT APPLICATION NUMBER: US/09/989,920
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/252,500
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 284
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 214
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-214

  Query Match             6.2%;  Score 5;  DB 9;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 LSSLL 10
           |||||
Db      20 LSSLL 24


RESULT 15
US-09-925-300-1735
; Sequence 1735, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT:  Steve Ruben
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA101
;  CURRENT APPLICATION NUMBER: US/09/925,300
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1735
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-300-1735

  Query Match             6.2%;  Score 5;  DB 10;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 4.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      43 RVPSP 47
           |||||
Db       4 RVPSP 8


Search completed: May 16, 2003, 15:36:37
Job time : 29 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:34:10 ; Search time 16 Seconds
(without alignments)
486.681 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 21742

Minimum DB seq length: 23
Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6	7.4	52	2	G84477	hypothetical prote
2	6	7.4	77	2	T49620	hypothetical prote
3	6	7.4	79	2	A82667	single-stranded DN
4	5	6.2	23	2	D33174	trimatin - wheat (
5	5	6.2	26	2	D49533	T-cell receptor be
6	5	6.2	26	2	C33174	avematin - oat (fr
7	5	6.2	26	4	I55277	hemoglobin alpha c
8	5	6.2	28	2	PH0231	T-cell receptor Vb
9	5	6.2	29	2	E47719	house-dust-mite-re
10	5	6.2	29	2	S58390	T-cell receptor be
11	5	6.2	31	2	PH0236	T-cell receptor Vb
12	5	6.2	33	2	I52083	major acute phase
13	5	6.2	35	2	PS0128	H-2 class I histoc
14	5	6.2	41	2	B81140	50S ribosomal prot
15	5	6.2	41	2	D82269	probable ribosomal
16	5	6.2	41	2	E82556	50S ribosomal prot
17	5	6.2	42	2	D30603	T-cell receptor be
18	5	6.2	44	2	S17573	thaumatin-like pro
19	5	6.2	47	2	AG0380	probable ribosomal
20	5	6.2	47	2	S70809	pilC-like protein
21	5	6.2	49	2	S13077	hypothetical prote
22	5	6.2	51	2	G84217	hypothetical prote
23	5	6.2	54	2	T42306	hypothetical prote
24	5	6.2	59	2	S01877	NADH2 dehydrogenas
25	5	6.2	59	4	T07378	probable hydroxyme
26	5	6.2	61	2	C30603	T-cell receptor be
27	5	6.2	61	2	AB1587	transcription regu
28	5	6.2	61	2	C69260	hypothetical prote
29	5	6.2	61	2	S21942	surface lectin, 17
30	5	6.2	61	2	AI3331	cbiM protein [impo
31	5	6.2	62	2	S30239	hypothetical prote
32	5	6.2	62	2	H81337	probable periplasm
33	5	6.2	63	2	D82547	hypothetical prote
34	5	6.2	63	2	H89761	hypothetical prote
35	5	6.2	63	2	AC2553	hypothetical prote
36	5	6.2	64	2	E81078	hypothetical prote
37	5	6.2	64	2	S01103	hypothetical prote
38	5	6.2	66	2	T12638	homeotic protein H
39	5	6.2	66	2	D72580	hypothetical prote
40	5	6.2	68	2	D27578	T-cell receptor be
41	5	6.2	68	2	S52778	T-cell receptor be
42	5	6.2	68	2	A88950	protein R09B5.10 [
43	5	6.2	68	2	H85826	hypothetical prote
44	5	6.2	68	2	AG3335	hypothetical prote
45	5	6.2	69	2	G95002	hypothetical prote

ALIGNMENTS

RESULT 1
G84477
hypothetical protein At2g06480 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: G84477
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y., M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84477
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-52 <STO>
A;Cross-references: GB:AE002093; NID:g4584353; PIDN:AAD25147.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g06480
A;Map position: 2

Query Match 7.4%; Score 6; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      74 VPGALP 79
           ||||||
Db       2 VPGALP 7
```

RESULT 2
T49620
hypothetical protein B5O22.20 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49620
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49620
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <SCH>
A;Cross-references: EMBL:AL355932; GSPDB:GN00116; NCSP:B5O22.20
A;Experimental source: BAC clone B5O22; strain OR74A
C;Genetics:
A;Gene: NCSP:B5O22.20
A;Map position: 6
A;Introns: 35/2

Query Match 7.4%; Score 6; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 77;

```
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 LCILLL 15
           ||||||
Db      62 LCILLL 67


RESULT 3
A82667
single-stranded DNA binding protein XF1558 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
C;Accession: A82667
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82667
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-79 <SIM>
A;Cross-references: GB:AE003985; GB:AE003849; NID:g9106593; PIDN:AAF84367.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1558
C;Superfamily: bacterial single-stranded DNA-binding protein; single-stranded DNA-bindin

  Query Match             7.4%;  Score 6;  DB 2;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 78;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      26 RRPAKA 31
           ||||||
Db      53 RRPAKA 58


RESULT 4
D33174
trimatin - wheat (fragment)
C;Species: Triticum aestivum (common wheat)
C;Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 12-Apr-1995
C;Accession: D33174
R;Vigers, A.J.; Roberts, W.K.; Selitrennikoff, C.P.
submitted to the Protein Sequence Database, May 1991
A;Reference number: A33174
A;Accession: D33174
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-23 <VIG>
C;Superfamily: thaumatin I

  Query Match             6.2%;  Score 5;  DB 2;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 3.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      75 PGALP 79
           |||||
Db      15 PGALP 19


RESULT 5
D49533
T-cell receptor beta chain (V-D-J region, clone HD8, LCMV-specific) - mouse (fragment)
N;Alternate names: lymphocytic choriomeningitis virus-specific T-cell receptor beta cha
C;Species: Mus musculus (house mouse)
C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
C;Accession: D49533
R;Horwitz, M.S.; Yanagi, Y.; Oldstone, M.B.
J. Virol. 68, 352-357, 1994
A;Title: T-cell receptors from virus-specific cytotoxic T lymphocytes recognizing a sin
A;Reference number: A49533; MUID:94076427; PMID:7504738
A;Accession: D49533
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-26 <HOR>
A;Experimental source: BALB/c, cytotoxic T lymphocytes
A;Note: sequence extracted from NCBI backbone (NCBIN:142930, NCBIP:142931)
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

  Query Match             6.2%;  Score 5;  DB 2;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RLLVL 6
           |||||
Db      19 RLLVL 23


RESULT 6
C33174
avematin - oat (fragment)
C;Species: Avena sativa (oat)
C;Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 12-Apr-1995
C;Accession: C33174
R;Vigers, A.J.; Roberts, W.K.; Selitrennikoff, C.P.
submitted to the Protein Sequence Database, May 1991
A;Reference number: A33174
A;Accession: C33174
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-26 <VIG>
C;Superfamily: thaumatin I

  Query Match             6.2%;  Score 5;  DB 2;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      75 PGALP 79
           |||||
Db      15 PGALP 19


RESULT 7
I55277
hemoglobin alpha chain thalassemia mutant Brooklyn - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 31-May-1996 #text_change 20-Apr-2000
C;Accession: I55277
R;Safaya, S.; Rieder, R.F.
J. Biol. Chem. 263, 4328-4332, 1988
A;Title: Dysfunctional alpha-globin gene in hemoglobin H disease in blacks. A dinucleot
A;Reference number: I55277; MUID:88153756; PMID:2831226
A;Accession: I55277
A;Molecule type: DNA
A;Residues: 1-26 <SAF>
A;Cross-references: GB:M22814; NID:g183799; PIDN:AAA52631.1; PID:g183801
A;Note: GenBank entry HUMHBADY2 repeats the normal sequence, which was not redetermine
C;Genetics:
A;Gene: GDB:HBA2
A;Cross-references: GDB:119294; OMIM:141850
```

```
A;Map position: 16p13.3-16p13.3

  Query Match              6.2%;  Score 5;  DB 4;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      67 PEPRL 71
           |||||
Db      19 PEPRL 23


RESULT 8
PH0231
T-cell receptor Vb CDR3, carrier PBL Vb 6.sbt - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 30-May-1997
C;Accession: PH0231
R;Hara, H.; Morita, M.; Iwaki, T.; Hatae, T.; Itoyama, Y.; Kitamoto, T.; Akizuki, S.; Go
submitted to JIPID, June 1994
A;Description: Detection of HTLV-I proviral DNA and analysis of T cell receptor Vb CDR3
A;Reference number: PH0227
A;Accession: PH0231
A;Molecule type: mRNA
A;Residues: 1-28 <HAR>
A;Experimental source: spinal cord
C;Genetics:
A;Map position: 7
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: receptor

  Query Match              6.2%;  Score 5;  DB 2;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RLLVL 6
           |||||
Db      19 RLLVL 23


RESULT 9
E47719
house-dust-mite-reactive T-cell receptor beta chain (CD4+ clone DE26, V(D)J junctional r
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
C;Accession: E47719
R;Wedderburn, L.R.; O'Hehir, R.E.; Hewitt, C.R.; Lamb, J.R.; Owen, M.J.
Proc. Natl. Acad. Sci. U.S.A. 90, 8214-8218, 1993
A;Title: In vivo clonal dominance and limited T-cell receptor usage in human CD4+ T-cell
A;Reference number: A47719; MUID:93376774; PMID:8367485
A;Accession: E47719
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-29 <WED>
A;Note: sequence extracted from NCBI backbone (NCBIN:137825, NCBIP:137831)
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

  Query Match              6.2%;  Score 5;  DB 2;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 3.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RLLVL 6
           |||||
Db      21 RLLVL 25


RESULT 10
S58390
T-cell receptor beta-chain Vb6-Jb2.5 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 29-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 23-Jul-1999
C;Accession: S58390
R;Johnston, S.L.; Strausbauch, M.; Sarkar, G.; Wettstein, P.J.
Nucleic Acids Res. 23, 3074-3075, 1995
A;Title: A novel method for sequencing members of multi-gene families.
A;Reference number: S58384; MUID:95388532; PMID:7659534
A;Accession: S58390
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-29 <JOH>
A;Cross-references: EMBL:U20304; NID:g663131; PIDN:AAA62251.1; PID:g663132
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
A;Note: only a part of the coding sequence is given
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

  Query Match              6.2%;  Score 5;  DB 2;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 3.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RLLVL 6
           |||||
Db      25 RLLVL 29


RESULT 11
PH0236
T-cell receptor Vb CDR3, Ctr2 TCR Vb12 CDR 3aa.sbt - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 09-May-1997
C;Accession: PH0236
R;Hara, H.; Morita, M.; Iwaki, T.; Hatae, T.; Itoyama, Y.; Kitamoto, T.; Akizuki, S.; C
submitted to JIPID, June 1994
A;Description: Detection of HTLV-I proviral DNA and analysis of T cell receptor Vb CDR3
A;Reference number: PH0227
A;Accession: PH0236
A;Molecule type: mRNA
A;Residues: 1-31 <HAR>
A;Experimental source: spinal cord
C;Genetics:
A;Map position: 7
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: receptor

  Query Match              6.2%;  Score 5;  DB 2;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RLLVL 6
           |||||
Db      22 RLLVL 26


RESULT 12
I52083
major acute phase protein - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999
C;Accession: I52083
R;Anderson, K.P.; Martin, A.D.; Heath, E.C.
Arch. Biochem. Biophys. 233, 624-635, 1984
A;Title: Rat major acute-phase protein: Biosynthesis and characterization of a cDNA clo
A;Reference number: I52083; MUID:85021400; PMID:6207775
A;Accession: I52083
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-33 <RES>
A;Cross-references: GB:M26758; NID:g205301; PIDN:AAA41567.1; PID:g205302
C;Superfamily: kininogen; cystatin homology

  Query Match              6.2%;  Score 5;  DB 2;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 ILLLC 16
           |||||
```

```
Db      2 ILLLC 6


RESULT 13
PS0128
H-2 class I histocompatibility antigen T10 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 12-Apr-1995
C;Accession: PS0128
R;Brorson, K.A.; Hunt III, S.W.; Hunkapiller, T.; Sun, Y.H.; Cheroutre, H.; Nickerson, D
J. Exp. Med. 170, 1837-1858, 1989
A;Title: Comparison of exon 5 sequences from 35 class I genes of the BALB/c mouse.
A;Reference number: PS0114; MUID:90063453; PMID:2584927
A;Accession: PS0128
A;Molecule type: DNA
A;Residues: 1-35 <BRO>
A;Experimental source: strain BALB/c
A;Note: this sequence is encoded by exon 5
C;Comment: This protein is a surface glycoprotein noncovalently associated with beta 2-m
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: glycoprotein; transmembrane protein; transplantation antigen
F;1-9/Domain: connecting peptide #status predicted <COP>
F;10-30/Domain: transmembrane #status predicted <TRS>
F;31-35/Domain: intracellular #status predicted <INT>

  Query Match             6.2%;  Score 5;  DB 2;  Length 35;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      73 VVPGA 77
           |||||
Db      16 VVPGA 20


RESULT 14
B81140
50S ribosomal protein L36 NMB0941 [imported] - Neisseria meningitidis (strain MC58 serog
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001
C;Accession: B81140; D81880
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: B81140
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-41 <TET>
A;Cross-references: GB:AE002445; GB:AE002098; NID:g7226173; PIDN:AAF41347.1; PID:g722617
A;Experimental source: serogroup B, strain MC58
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: D81880
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-41 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84399.1; PID:g737983
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: rpmJ2; NMB0941; NMA1137
C;Superfamily: Escherichia coli ribosomal protein L36

  Query Match             6.2%;  Score 5;  DB 2;  Length 41;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 VLSSL 9
            |||||
Db       3 VLSSL 7


RESULT 15
D82269
probable ribosomal protein L36 VC0879 [imported] - Vibrio cholerae (strain N16961 serog
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: D82269
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: D82269
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-41 <HEI>
A;Cross-references: GB:AE004172; GB:AE003852; NID:g9655323; PIDN:AAF94041.1; GSPDB:GN00
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0879
A;Map position: 1
C;Superfamily: Escherichia coli ribosomal protein L36

  Query Match             6.2%;  Score 5;  DB 2;  Length 41;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 VLSSL 9
            |||||
Db       3 VLSSL 7


Search completed: May 16, 2003, 15:35:27
Job time : 18 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:30:39 ; Search time 12 Seconds
(without alignments)
279.965 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7400

Minimum DB seq length: 23
Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	6	7.4	23	1	PRP1_RAT	P10165	rattus norv
2	5	6.2	44	1	THHR_HORVU	P33044	hordeum vul
3	5	6.2	48	1	ECNB_CITFR	P56550	citrobacter
4	5	6.2	59	1	NU5M_ARTSA	P19047	artemia sal
5	5	6.2	61	1	Y083_ARCFU	O30153	archaeoglob
6	5	6.2	66	1	CXA2_CONTE	Q9xzk7	conus texti
7	5	6.2	67	1	VC18_SPVKA	P32217	swinepox vi
8	5	6.2	68	1	L2MU_ADECC	Q65953	canine aden
9	5	6.2	70	1	RS21_ECOLI	P02379	escherichia
10	5	6.2	74	1	YC47_ODOSI	P49542	odontella s
11	5	6.2	74	1	YC55_HAEIN	P44141	haemophilus
12	5	6.2	80	1	GRFA_SFVKA	P08441	shope fibro
13	5	6.2	80	1	ICTA_SYNY3	P27372	synechocyst
14	4	4.9	23	1	ACHG_ELEEL	P09692	electrophor
15	4	4.9	23	1	UHA4_HUMAN	P49289	homo sapien
16	4	4.9	24	1	REL1_ECOLI	P56980	escherichia
17	4	4.9	25	1	CR14_LITGI	P56229	litoria gil
18	4	4.9	25	1	SPIG_PSEUS	P82357	pseudacanth
19	4	4.9	26	1	RSHX_THETH	P32193	thermus the
20	4	4.9	27	1	CXM3_CONPU	P58925	conus purpu
21	4	4.9	28	1	Y16P_BPT4	P39248	bacteriopha
22	4	4.9	29	1	P2SM_LOXIN	P83046	loxosceles
23	4	4.9	29	1	PSAK_SPIOL	P14627	spinacia ol
24	4	4.9	30	1	PRT1_CLUPA	P02335	clupea pall
25	4	4.9	30	1	UDDP_SULAC	P80143	sulfolobus
26	4	4.9	31	1	DEF2_MESAU	P81466	mesocricetu
27	4	4.9	31	1	ER29_BOVIN	P81623	bos taurus
28	4	4.9	31	1	PETM_CYACA	Q9tlr5	cyanidium c
29	4	4.9	31	1	PSBT_CHLRE	P37256	chlamydomon
30	4	4.9	32	1	NEUB_PIG	P01297	sus scrofa
31	4	4.9	32	1	PRT_ORYLA	Q91185	oryzias lat
32	4	4.9	32	1	TX29_PHONI	P29426	phoneutria
33	4	4.9	33	1	DEF1_MESAU	P81465	mesocricetu
34	4	4.9	33	1	DEF4_MESAU	P81468	mesocricetu
35	4	4.9	33	1	PEN3_ADECU	P35987	canine aden
36	4	4.9	33	1	Y50A_MYCTU	Q9cb56	mycobacteri
37	4	4.9	34	1	PRT_DICLA	Q9ps27	dicentrarch
38	4	4.9	34	1	PRT_PERFV	P29629	perca flave
39	4	4.9	35	1	PSPC_PIG	P15785	sus scrofa
40	4	4.9	35	1	SCX1_BUTSI	P15229	buthus sind
41	4	4.9	35	1	Y37_BPT3	P20325	bacteriopha
42	4	4.9	35	1	Y845_BORBU	O51785	borrelia bu
43	4	4.9	36	1	MYPC_RAT	P56741	rattus norv
44	4	4.9	36	1	PGKH_CHLFU	P36232	chlorella f
45	4	4.9	36	1	PSBY_ODOSI	P49543	odontella s

ALIGNMENTS

RESULT 1
PRP1_RAT
ID PRP1_RAT STANDARD; PRT; 23 AA.
AC P10165;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acidic proline-rich protein PRP18 precursor (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Parotid gland;
RX MEDLINE=86033799; PubMed=3840480;
RA Clements S., Mehansho H., Carlson D.M.;
RT "Novel multigene families encoding highly repetitive peptide
RT sequences. Sequence analyses of rat and mouse proline-rich protein
RT cDNAs.";
RL J. Biol. Chem. 260:13471-13477(1985).
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; M11899; AAA41956.1; -.
KW Repeat; Parotid gland; Multigene family; Saliva; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 >23 ACIDIC PROLINE-RICH PROTEIN PRP18.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2380 MW; 875B4F61FD056949 CRC64;

Query Match 7.4%; Score 6; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LLVLSS 8
||||||
Db 10 LLVLSS 15

RESULT 2
THHR_HORVU
ID THHR_HORVU STANDARD; PRT; 44 AA.
AC P33044;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antifungal protein R (Fragment).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. BOMI RISO 1508;
RX   MEDLINE=92037994; PubMed=1936240;
RA   Hejgaard J., Jacobsen S., Svendsen I.;
RT   "Two antifungal thaumatin-like proteins from barley grain.";
RL   FEBS Lett. 291:127-131(1991).
CC   -!- FUNCTION: HAS ANTIFUNGAL ACTIVITY. INHIBITS THE GROWTH OF
CC       TRICHODERMA VIRIDAE AND CANDIDA ALBICANS.
CC   -!- SIMILARITY: BELONGS TO THE THAUMATIN FAMILY.
DR   PIR; S17573; S17573.
DR   HSSP; P25871; 1AUN.
DR   InterPro; IPR001938; Thaumatin.
DR   Pfam; PF00314; thaumatin; 1.
DR   ProDom; PD001321; Thaumatin; 1.
DR   PROSITE; PS00316; THAUMATIN; PARTIAL.
KW   Plant defense; Fungicide.
FT   NON_TER      44     44
SQ   SEQUENCE   44 AA;  4453 MW;  920713BCE551C1E4 CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      75 PGALP 79
           |||||
Db      15 PGALP 19


RESULT 3
ECNB_CITFR
ID   ECNB_CITFR     STANDARD;      PRT;    48 AA.
AC   P56550;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Entericidin B precursor.
GN   ECNB.
OS   Citrobacter freundii.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Citrobacter.
OX   NCBI_TaxID=546;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OS60;
RX   MEDLINE=98344100; PubMed=9677290;
RA   Bishop R.E., Leskiw B.K., Hodges R.S., Kay C.M., Weiner J.H.;
RT   "The entericidin locus of Escherichia coli and its implications for
RT   programmed bacterial cell death.";
RL   J. Mol. Biol. 280:583-596(1998).
CC   -!- FUNCTION: PLAYS A ROLE IN THE BACTERIOLYSIS. IS ACTIVATED UNDER
CC       CONDITIONS OF HIGH OSMOLARITY BY THE FACTOR SIGMA S. ENTERICIDIN A
CC       FUNCTIONS AS A ANTIDOTE.
CC   -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor.
CC   -!- SIMILARITY: STRONG, TO ECNA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U21727; AAC46459.1; -.
KW   Membrane; Lipoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22     48       ENTERICIDIN B.
FT   LIPID        22     22       N-ACYL DIGLYCERIDE (PROBABLE).
SQ   SEQUENCE   48 AA;  4767 MW;  D951A3C8236FD18E CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 48;
  Best Local Similarity   100.0%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 LVLSS 8
            |||||
Db      13 LVLSS 17


RESULT 4
NU5M_ARTSA
ID   NU5M_ARTSA     STANDARD;      PRT;    59 AA.
AC   P19047;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3) (Fragment).
GN   ND5.
OS   Artemia salina (Brine shrimp).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC   Branchiopoda; Anostraca; Artemiidae; Artemia.
OX   NCBI_TaxID=85549;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88289417; PubMed=3135541;
RA   Batuecas B., Garesse R., Calleja M., Valverde J.R., Marco R.;
RT   "Genome organization of Artemia mitochondrial DNA.";
RL   Nucleic Acids Res. 16:6515-6529(1988).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X07663; CAA30510.1; -.
DR   PIR; S01877; S01877.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   59 AA;  6585 MW;  60C9E6E04827DC23 CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 LSSLL 10
            |||||
Db      46 LSSLL 50


RESULT 5
Y083_ARCFU
ID   Y083_ARCFU     STANDARD;      PRT;    61 AA.
AC   O30153;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein AF0083.
GN   AF0083.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC   Archaeoglobaceae; Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
```

```
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001100; AAB91149.1; -.
DR   TIGR; AF0083; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   61 AA;  7053 MW;  4096E1B2A28E7D36 CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 61;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 RPAKA 31
           |||||
Db      43 RPAKA 47


RESULT 6
CXA2_CONTE
ID   CXA2_CONTE     STANDARD;      PRT;    66 AA.
AC   Q9XZK7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Alpha-type conotoxin Tx2 precursor.
OS   Conus textile (Cloth-of-gold cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=6494;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom duct;
RX   MEDLINE=20037955; PubMed=10573284;
RA   Lu B.-S., Yu F., Zhao D., Huang P.-T., Huang C.-F.;
RT   "Conopeptides from Conus striatus and Conus textile by cDNA
RT   cloning.";
RL   Peptides 20:1139-1144(1999).
CC   -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they
CC       bind to the nicotinic acetylcholine receptors (nAChR) and thus
CC       inhibit them (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF146353; AAD31913.1; -.
KW   Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW   Acetylcholine receptor inhibitor; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     48       POTENTIAL.
FT   PEPTIDE      49     66       ALPHA-TYPE CONOTOXIN TX2.
FT   DISULFID     51     57       BY SIMILARITY.
FT   DISULFID     52     65       BY SIMILARITY.
SQ   SEQUENCE   66 AA;  7254 MW;  EDDB59BBAB94F26F CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 66;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      33 SGRRT 37
           |||||
Db      24 SGRRT 28


RESULT 7
VC18_SPVKA
ID   VC18_SPVKA     STANDARD;      PRT;    67 AA.
AC   P32217;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Hypothetical protein C18.
GN   C18L.
OS   Swinepox virus (strain Kasza) (SPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Suipoxvirus.
OX   NCBI_TaxID=10277;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94069924; PubMed=8249275;
RA   Massung R.F., Jayarama V., Moyer R.W.;
RT   "DNA sequence analysis of conserved and unique regions of swinepox
RT   virus: identification of genetic elements supporting phenotypic
RT   observations including a novel G protein-coupled receptor
RT   homologue.";
RL   Virology 197:511-528(1993).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L22013; AAC37853.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   67 AA;  8258 MW;  4D34A27408B74975 CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 LCILL 14
           |||||
Db      16 LCILL 20


RESULT 8
L2MU_ADECC
ID   L2MU_ADECC     STANDARD;      PRT;    68 AA.
AC   Q65953;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Late L2 mu core protein precursor (pMu) (Protein X).
GN   PX.
OS   Canine adenovirus type 1 (strain CLL), and
OS   Canine adenovirus type 1 (strain RI261).
```

```
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=69150, 69151;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLL;
RA   Campbell J.B., Zhao Y.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RI261;
RX   MEDLINE=97275900; PubMed=9129661;
RA   Morrison M.D., Onions D.E., Nicolson L.;
RT   "Complete DNA sequence of canine adenovirus type 1.";
RL   J. Gen. Virol. 78:873-878(1997).
CC   -!- FUNCTION: THE ROLE OF THE PRECURSOR MIGHT BE TO CONDENSE THE
CC       VIRAL PROCHROMATIN FOR ENCAPSIDATION BY VIRTUE OF THE TWO
CC       BASIC DOMAINS (BY SIMILARITY).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U55001; AAB05441.1; -.
DR   EMBL; Y07760; CAA69064.1; -.
KW   Core protein; DNA-binding; Late protein.
FT   PROPEP          1     21       POTENTIAL.
FT   CHAIN          22     39       LATE L2 MU CORE PROTEIN.
FT   PROPEP         40     68       POTENTIAL.
FT   SITE           21     22       CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT                                  (POTENTIAL).
FT   SITE           39     40       CLEAVAGE (BY ADENOVIRUS PROTEASE)
FT                                  (POTENTIAL).
SQ   SEQUENCE   68 AA;  7226 MW;  794038F721AB152D CRC64;

  Query Match              6.2%;  Score 5;  DB 1;  Length 68;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      34 GRRTR 38
           |||||
Db      26 GRRTR 30


RESULT 9
RS21_ECOLI
ID   RS21_ECOLI     STANDARD;      PRT;    70 AA.
AC   P02379; Q8ZI69;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S21.
GN   RPSU OR B3065 OR Z4418 OR ECS3948 OR STM3209 OR STY3388 OR YPO0645.
OS   Escherichia coli,
OS   Escherichia coli O157:H7,
OS   Salmonella typhimurium,
OS   Salmonella typhi, and
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562, 83334, 602, 601, 632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli;
RX   MEDLINE=83129424; PubMed=6186393;
RA   Burton Z.F., Gross C.A., Watanabe K.K., Burgess R.R.;
RT   "The operon that encodes the sigma subunit of RNA polymerase also
RT   encodes ribosomal protein S21 and DNA primase in E. coli K12.";
RL   Cell 32:335-349(1983).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=K12;
RX   MEDLINE=83218520; PubMed=6222240;
RA   Lupski J.R., Smiley B.L., Godson G.N.;
RT   "Regulation of the rpsU-dnaG-rpoD macromolecular synthesis operon and
RT   the initiation of DNA replication in Escherichia coli K-12.";
RL   Mol. Gen. Genet. 189:48-57(1983).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
RN   [6]
RP   SEQUENCE.
RC   SPECIES=E.coli; STRAIN=K;
RX   MEDLINE=76119561; PubMed=765257;
RA   Vandekerckhove J., Rombauts W., Peeters B., Wittmann-Liebold B.;
RT   "Determination of the complete amino acid sequence of protein S21
RT   from Escherichia coli ribosomes.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 356:1955-1976(1975).
RN   [7]
RP   SEQUENCE OF 25-34, AND CROSS-LINKING TO RRNA.
RC   SPECIES=E.coli; STRAIN=MRE-600;
RX   PubMed=7556101;
RA   Urlaub H., Kruft V., Bischof O., Mueller E.-C., Wittmann-Liebold B.;
RT   "Protein-rRNA binding features and their structural and functional
RT   implications in ribosomes as determined by cross-linking studies.";
RL   EMBO J. 14:4578-4588(1995).
RN   [8]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.typhimurium;
RX   MEDLINE=86137422; PubMed=3005129;
RA   Erickson B.D., Burton Z.F., Watanabe K.K., Burgess R.R.;
RT   "Nucleotide sequence of the rpsU-dnaG-rpoD operon from Salmonella
RT   typhimurium and a comparison of this sequence with the homologous
RT   operon of Escherichia coli.";
RL   Gene 40:67-78(1985).
RN   [9]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
```

```
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [10]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.typhi; STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [11]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Y.pestis; STRAIN=CO-92 / Biovar Orientalis;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
RN   [12]
RP   MASS SPECTROMETRY.
RC   SPECIES=E.coli; STRAIN=K12;
RX   MEDLINE=99196679; PubMed=10094780;
RA   Arnold R.J., Reilly J.P.;
RT   "Observation of Escherichia coli ribosomal proteins and their
RT   posttranslational modifications by mass spectrometry.";
RL   Anal. Biochem. 269:105-112(1999).
CC   -!- MASS SPECTROMETRY: MW=8368.8; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J01687; AAA24599.1; -.
DR   EMBL; V00346; CAA23635.1; -.
DR   EMBL; U28379; AAA89145.1; -.
DR   EMBL; AE000388; AAC76101.1; -.
DR   EMBL; AE005535; AAG58199.1; -.
DR   EMBL; AP002564; BAB37371.1; -.
DR   EMBL; M14427; AAA27240.1; -.
DR   EMBL; AE008847; AAL22083.1; -.
DR   EMBL; AL627278; CAD07734.1; -.
DR   EMBL; AJ414144; CAC89499.1; -.
DR   PIR; A02749; R3EC21.
DR   PIR; A23985; R3EB21.
DR   EcoGene; EG10920; rpsU.
DR   StyGene; SG10376; rpsU.
DR   InterPro; IPR001911; Ribosomal_S21.
DR   Pfam; PF01165; Ribosomal_S21; 1.
DR   PRINTS; PR00976; RIBOSOMALS21.
DR   ProDom; PD005521; Ribosomal_S21; 1.
DR   TIGRFAMs; TIGR00030; S21p; 1.
DR   PROSITE; PS01181; RIBOSOMAL_S21; 1.
KW   Ribosomal protein; Complete proteome.
FT   INIT_MET      0      0
SQ   SEQUENCE   70 AA;  8369 MW;  2BB7C7B2FBB29013 CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      35 RRTRL 39
           |||-||
Db      65 RRTRL 69


RESULT 10
YC47_ODOSI
ID   YC47_ODOSI     STANDARD;      PRT;    74 AA.
AC   P49542;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Hypothetical 8.4 kDa protein ycf47 (ORF74).
GN   YCF47.
OS   Odontella sinensis (Marine centric diatom).
OG   Chloroplast.
OC   Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC   Biddulphiophycidae; Eupodiscales; Eupodiscaceae; Odontella.
OX   NCBI_TaxID=2839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT   "The chloroplast genome of a chlorophyll a+c-containing alga,
RT   Odontella sinensis.";
RL   Plant Mol. Biol. Rep. 13:336-342(1995).
CC   -!- SIMILARITY: BELONGS TO THE YCF47 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z67753; CAA91678.1; -.
KW   Chloroplast; Hypothetical protein.
SQ   SEQUENCE   74 AA;  8436 MW;  A80A858F7A939C72 CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 SPNST 50
           |||||
Db      41 SPNST 45


RESULT 11
YC55_HAEIN
ID   YC55_HAEIN     STANDARD;      PRT;    74 AA.
AC   P44141;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hypothetical protein HI1255.
GN   HI1255.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
```

```
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32805; AAC22910.1; -.
DR   TIGR; HI1255; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   74 AA;  8446 MW;  CC22E0A66C0C98D3 CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      24 GKRRP 28
           |||||
Db      13 GKRRP 17


RESULT 12
GRFA_SFVKA
ID   GRFA_SFVKA     STANDARD;      PRT;    80 AA.
AC   P08441;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Growth factor.
OS   Shope fibroma virus (strain Kasza) (SFV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Leporipoxvirus.
OX   NCBI_TaxID=10272;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87172751; PubMed=3031480;
RA   Chang W., Upton C., Hu S.-L., Purchio A.F., McFadden G.;
RT   "The genome of Shope fibroma virus, a tumorigenic poxvirus, contains
RT   a growth factor gene with sequence similarity to those encoding
RT   epidermal growth factor and transforming growth factor alpha.";
RL   Mol. Cell. Biol. 7:535-540(1987).
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M15921; AAA66873.1; -.
DR   PIR; A26723; EGVZSF.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001336; EGF_1.
DR   Pfam; PF00008; EGF; 1.
DR   PRINTS; PR00009; EGFTGF.
DR   SMART; SM00181; EGF; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; FALSE_NEG.
KW   EGF-like domain; Growth factor; Glycoprotein.
FT   DOMAIN       29     73       EGF-LIKE.
FT   DISULFID     33     47       BY SIMILARITY.
FT   DISULFID     41     61       BY SIMILARITY.
FT   DISULFID     63     72       BY SIMILARITY.
FT   CARBOHYD     44     44       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     54     54       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   80 AA;  9210 MW;  C48D30E878D2ED58 CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 3.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 SLLCI 12
             |||||
Db       9 SLLCI 13


RESULT 13
ICTA_SYNY3
ID   ICTA_SYNY3     STANDARD;      PRT;    80 AA.
AC   P27372;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Inorganic carbon transport protein.
GN   ICTA OR SSR1386.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ogawa T.;
RT   "Cloning and inactivation of a gene essential to inorganic carbon
RT   transport of Synechocystis PCC6803.";
RL   Plant Physiol. 96:280-284(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA   Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA   Yamada M., Yasuda M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M73833; AAA27290.1; -.
DR   EMBL; D90911; BAA18123.1; -.
DR   PIR; JQ1959; JQ1959.
KW   Transport; Complete proteome.
SQ   SEQUENCE   80 AA;  9252 MW;  311AD10FA3B6136D CRC64;

  Query Match             6.2%;  Score 5;  DB 1;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 3.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLVLS 7
             |||||
Db      64 LLVLS 68


RESULT 14
ACHG_ELEEL
```

```
ID   ACHG_ELEEL     STANDARD;      PRT;    23 AA.
AC   P09692;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Acetylcholine receptor protein, gamma chain (Fragment).
OS   Electrophorus electricus (Electric eel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Gymnotiformes;
OC   Electrophoridae; Electrophorus.
OX   NCBI_TaxID=8005;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=83065199; PubMed=6959131;
RA   Conti-Tronconi B.M., Hunkapiller M.W., Lindstrom J.M., Raftery M.A.;
RT   "Subunit structure of the acetylcholine receptor from Electrophorus
RT   electricus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:6489-6493(1982).
CC   -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC       EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC       LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC       MEMBRANE.
CC   -!- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,
CC       DELTA, AND GAMMA CHAINS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR   PIR; C27262; C27262.
DR   InterPro; IPR001175; Neur_channel.
DR   Pfam; PF02931; Neur_chan_LBD; 1.
DR   PROSITE; PS00236; NEUROTR_ION_CHANNEL; PARTIAL.
KW   Receptor; Postsynaptic membrane; Ionic channel; Transmembrane.
FT   NON_TER       23      23
SQ   SEQUENCE   23 AA;  2717 MW;  DA88F7E9900C024B CRC64;

  Query Match             4.9%;  Score 4;  DB 1;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 RPAK 30
           ||||
Db      19 RPAK 22


RESULT 15
UHA4_HUMAN
ID   UHA4_HUMAN     STANDARD;      PRT;    23 AA.
AC   P49289;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Unknown protein from 2D-page of heart (Fragments).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Heart;
RX   MEDLINE=96007936; PubMed=7498159;
RA   Kovalyov L.I., Shishkin S.S., Efimochkin A.S., Kovalyova M.A.,
RA   Ershova E.S., Egorov T.A., Musalyamov A.K.;
RT   "The major protein expression profile and two-dimensional protein
RT   database of human heart.";
RL   Electrophoresis 16:1160-1169(1995).
CC   -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE MITOCHONDRIAL
CC       ACONITASE.
FT   NON_TER       1       1
FT   NON_CONS     16      17
FT   NON_TER      23      23
SQ   SEQUENCE   23 AA;  2388 MW;  90EF513CEEB8B451 CRC64;

  Query Match             4.9%;  Score 4;  DB 1;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      50 TNLK 53
           ||||
Db      20 TNLK 23


Search completed: May 16, 2003, 15:34:26
Job time : 14 secs
```

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:31:20 ; Search time 30 Seconds
(without alignments)
556.327 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 70278

Minimum DB seq length: 23
Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7	8.6	51	5	O61655	O61655 ceratitis c
2	7	8.6	73	16	Q8XRN8	Q8xrn8 ralstonia s
3	6	7.4	30	4	Q96D69	Q96d69 homo sapien
4	6	7.4	52	10	Q9SKI7	Q9ski7 arabidopsis
5	6	7.4	73	4	Q9BT96	Q9bt96 homo sapien
6	6	7.4	76	10	Q9XER1	Q9xer1 gossypium h
7	6	7.4	77	3	Q9P5V0	Q9p5v0 neurospora
8	6	7.4	78	16	Q9A1T5	Q9a1t5 streptococc
9	6	7.4	79	12	Q91LC1	Q91lc1 white spot
10	6	7.4	79	16	Q9PD22	Q9pd22 xylella fas
11	5	6.2	23	10	Q9S906	Q9s906 triticum ae
12	5	6.2	24	8	Q94XE2	Q94xe2 tectocoris
13	5	6.2	25	11	Q9QUX1	Q9qux1 rattus sp.
14	5	6.2	25	13	Q9PWR9	Q9pwr9 xiphophorus
15	5	6.2	26	10	Q9S907	Q9s907 avena sativ
16	5	6.2	27	12	Q9Q9F2	Q9q9f2 hepatitis c
17	5	6.2	27	12	Q9Q9E6	Q9q9e6 hepatitis c
18	5	6.2	27	12	Q9Q9E2	Q9q9e2 hepatitis c
19	5	6.2	27	12	Q9Q9D6	Q9q9d6 hepatitis c
20	5	6.2	28	4	O43804	O43804 homo sapien
21	5	6.2	28	8	O98676	O98676 ondinea pur
22	5	6.2	29	10	Q9S8C7	Q9s8c7 secale cere
23	5	6.2	30	7	Q31234	Q31234 mus musculu
24	5	6.2	30	11	Q9WVD6	Q9wvd6 mus musculu
25	5	6.2	31	2	Q47323	Q47323 escherichia
26	5	6.2	32	7	Q9BCT8	Q9bct8 homo sapien
27	5	6.2	32	7	Q95HL3	Q95hl3 homo sapien
28	5	6.2	32	7	Q8SNB9	Q8snb9 homo sapien
29	5	6.2	32	11	Q9R1C2	Q9r1c2 mus musculu
30	5	6.2	33	11	Q63329	Q63329 rattus norv
31	5	6.2	34	4	Q96J93	Q96j93 homo sapien
32	5	6.2	35	2	Q9R4S7	Q9r4s7 rhodobacter
33	5	6.2	35	4	Q9NQY7	Q9nqy7 homo sapien
34	5	6.2	35	4	Q9UGE7	Q9uge7 homo sapien
35	5	6.2	37	2	Q9EYZ8	Q9eyz8 photobacter
36	5	6.2	37	4	Q9NQN5	Q9nqn5 homo sapien
37	5	6.2	37	5	Q9BM33	Q9bm33 themiste al
38	5	6.2	37	5	Q9BM32	Q9bm32 themiste al
39	5	6.2	39	3	Q8TGR6	Q8tgr6 saccharomyc
40	5	6.2	39	12	Q88527	Q88527 turkey herp
41	5	6.2	39	16	Q8XTC6	Q8xtc6 ralstonia s
42	5	6.2	40	12	Q91K09	Q91k09 hepatitis c
43	5	6.2	41	15	Q9PXE3	Q9pxe3 chimpanzee
44	5	6.2	41	16	Q9PAQ5	Q9paq5 xylella fas
45	5	6.2	41	16	Q9KTM3	Q9ktm3 vibrio chol

ALIGNMENTS

RESULT 1
O61655
ID O61655 PRELIMINARY; PRT; 51 AA.
AC O61655;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cytochrome P450 monooxygenase (Fragment).
OS Ceratitis capitata (Mediterranean fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Tephritoidea; Tephritidae; Ceratitis.
OX NCBI_TaxID=7213;
RN [1]
RP SEQUENCE FROM N.A.
RA Danielson P.B., Foster J.L.M., Cooper S.K., Fogleman J.C.;
RT "Diversity of Expressed cytochrome P450 Genes in the Adult
RT Mediterranean Fruit Fly, Ceratitis capitata.";
RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF056474; AAC13307.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
KW Monooxygenase.
FT NON_TER 1 1
SQ SEQUENCE 51 AA; 5854 MW; 0CA1F3B75BDECE44 CRC64;

Query Match 8.6%; Score 7; DB 5; Length 51;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       4 LVLSSLL 10
           |||||||
Db      10 LVLSSLL 16
```

RESULT 2
Q8XRN8
ID Q8XRN8 PRELIMINARY; PRT; 73 AA.
AC Q8XRN8;

```
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein RSp0793.
GN   RSP0793 OR RS01914.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OG   Plasmid megaplasmid.
OC   Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC   Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646080; CAD17944.1; -.
KW   Plasmid; Hypothetical protein; Complete proteome.
SQ   SEQUENCE   73 AA;  8333 MW;  8BEFAAC58DAE173C CRC64;

  Query Match             8.6%;  Score 7;  DB 16;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      23 EGKRRPA 29
           |||||||
Db      29 EGKRRPA 35


RESULT 3
Q96D69
ID   Q96D69      PRELIMINARY;      PRT;    30 AA.
AC   Q96D69;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 3.1 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Strausberg R.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC012777; AAH12777.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   30 AA;  3108 MW;  C5EE62C1A5F8405C CRC64;

  Query Match             7.4%;  Score 6;  DB 4;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 42;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      73 VVPGAL 78
           ||||||
Db      15 VVPGAL 20


RESULT 4
Q9SKI7
ID   Q9SKI7      PRELIMINARY;      PRT;    52 AA.
AC   Q9SKI7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   At2g06480 protein.
GN   AT2G06480.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC006420; AAD25147.1; -.
SQ   SEQUENCE   52 AA;  5985 MW;  82F906B7FE4DCEE1 CRC64;

  Query Match             7.4%;  Score 6;  DB 10;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      74 VPGALP 79
           ||||||
Db       2 VPGALP 7


RESULT 5
Q9BT96
ID   Q9BT96      PRELIMINARY;      PRT;    73 AA.
AC   Q9BT96;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Similar to hypothetical protein FLJ10719.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKIN;
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC004277; AAH04277.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   73 AA;  8217 MW;  64A7140446E960DD CRC64;

  Query Match             7.4%;  Score 6;  DB 4;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLVLSS 8
           ||||||
Db      38 LLVLSS 43


RESULT 6
Q9XER1
ID   Q9XER1      PRELIMINARY;      PRT;    76 AA.
AC   Q9XER1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
```

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   FS18A precursor.
GN   FS18A.
OS   Gossypium hirsutum (Upland cotton).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Malvales; Malvaceae; Gossypium.
OX   NCBI_TaxID=3635;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. SIOKRA 1-4;
RA   Orford S.J., Carney T.J., Olesnicky N.S., Timmis J.N.;
RT   "Characterization of a cotton gene expressed late in fiber cell
RT   elongation.";
RL   Theor. Appl. Genet. 0:0-0(1999).
DR   EMBL; AF114254; AAD31394.1; -.
KW   Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27     76       PROCESSED FS18A.
SQ   SEQUENCE   76 AA;  7556 MW;  0C3FFB299AE7DB57 CRC64;

  Query Match             7.4%;  Score 6;  DB 10;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 LSSLLC 11
          ||||||
Db     68 LSSLLC 73


RESULT 7
Q9P5V0
ID   Q9P5V0      PRELIMINARY;      PRT;    77 AA.
AC   Q9P5V0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Conserved hypothetical protein.
GN   B5O22.20.
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   German Neurospora genome project;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL355932; CAB91421.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   77 AA;  8846 MW;  D6FD4B9CB299A7E3 CRC64;

  Query Match             7.4%;  Score 6;  DB 3;  Length 77;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     10 LCILLL 15
          ||||||
Db     62 LCILLL 67


RESULT 8
Q9A1T5
ID   Q9A1T5      PRELIMINARY;      PRT;    78 AA.
AC   Q9A1T5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical protein SPy0105.
GN   SPY0105.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR   EMBL; AE006481; AAK33223.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   78 AA;  8898 MW;  6EC41E2909801EE1 CRC64;

  Query Match             7.4%;  Score 6;  DB 16;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 97;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 LVLSSL 9
          ||||||
Db     13 LVLSSL 18


RESULT 9
Q91LC1
ID   Q91LC1      PRELIMINARY;      PRT;    79 AA.
AC   Q91LC1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   ORF124 (Wsv247) (WSSV302).
OS   White spot syndrome virus (WSSV).
OC   Viruses; unclassified viruses.
OX   NCBI_TaxID=92652;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21342572; PubMed=11448154;
RA   van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA   Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;
RT   "The white spot syndrome virus DNA genome sequence.";
RL   Virology 286:7-22(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA   Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21548311; PubMed=11689662;
RA   Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RT   "Complete genome sequence of the shrimp white spot bacilliform
RT   virus.";
RL   J. Virol. 75:11811-11820(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TAIWAN;
RX   MEDLINE=20517548; PubMed=11062040;
RA   Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
RA   Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
RT   "Identification and characterization of a shrimp white spot syndrome
RT   virus (WSSV) gene that encodes a novel chimeric polypeptide of
RT   cellular-type thymidine kinase and thymidylate kinase.";
RL   Virology 277:100-110(2000).
RN   [6]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=TAIWAN;
RX   MEDLINE=21844071; PubMed=11853398;
RA   Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
RA   Lo C.F., Kou G.H.;
RT   "Identification of a nucleocapsid protein (VP35) gene of shrimp white
RT   spot syndrome virus and characterization of the motif important for
RT   targeting VP35 to the nuclei of transfected insect cells.";
RL   Virology 293:44-53(2002).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TAIWAN;
RA   Lo C.-F., Kou G.-H.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF369029; AAK77793.1; -.
DR   EMBL; AF332093; AAL33250.1; -.
DR   EMBL; AF440570; AAL89170.1; -.
SQ   SEQUENCE   79 AA;  9285 MW;  734EE8DBB60FDA86 CRC64;

  Query Match             7.4%;  Score 6;  DB 12;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 VLSSLL 10
           ||||||
Db      58 VLSSLL 63


RESULT 10
Q9PD22
ID   Q9PD22      PRELIMINARY;      PRT;    79 AA.
AC   Q9PD22;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Single-stranded DNA binding protein.
GN   XF1558.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9A5C;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE003985; AAF84367.1; -.
DR   HSSP; P02339; 1EYG.
DR   InterPro; IPR000424; SSB_protein.
DR   Pfam; PF00436; SSB; 1.
KW   Complete proteome.
SQ   SEQUENCE   79 AA;  8935 MW;  2DEEA09C35D4EADD CRC64;

  Query Match             7.4%;  Score 6;  DB 16;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      26 RRPAKA 31
           ||||||
Db      53 RRPAKA 58


RESULT 11
Q9S906
ID   Q9S906      PRELIMINARY;      PRT;    23 AA.
AC   Q9S906;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Antifungal protein (Fragment).
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92190627; PubMed=1799695;
RA   Vigers A.J., Roberts W.K., Selitrennikoff C.P.;
RT   "A new family of plant antifungal proteins.";
RL   Mol. Plant Microbe Interact. 4:315-323(1991).
DR   InterPro; IPR001938; Thaumatin.
DR   Pfam; PF00314; thaumatin; 1.
FT   NON_TER       1      1
FT   NON_TER      23     23
SQ   SEQUENCE   23 AA;  2291 MW;  C3D88E4A722C21D7 CRC64;

  Query Match             6.2%;  Score 5;  DB 10;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      75 PGALP 79
           |||||
Db      15 PGALP 19


RESULT 12
Q94XE2
ID   Q94XE2      PRELIMINARY;      PRT;    24 AA.
AC   Q94XE2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NADH dehydrogenase subunit 4 (Fragment).
GN   NAD4.
OS   Tectocoris diophthalmus (cotton harlequin bug).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC   Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomoidea;
OC   Pentatomidae; Tectocoris.
OX   NCBI_TaxID=159956;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21396409; PubMed=11504862;
RA   Shao R., Campbell N.J.H., Schmidt E.R., Barker S.C.;
RT   "Increased rate of gene rearrangement in the mitochondrial genomes of
RT   three orders of hemipteroid insects.";
RL   Mol. Biol. Evol. 18:1828-1832(2001).
```

```
DR   EMBL; AF335991; AAK55287.1; -.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   24 AA;  2871 MW;  972B55027DC29675 CRC64;

  Query Match             6.2%;  Score 5;  DB 8;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 FSIFS 21
           |||||
Db       2 FSIFS 6


RESULT 13
Q9QUX1
ID   Q9QUX1      PRELIMINARY;      PRT;    25 AA.
AC   Q9QUX1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   N-acetyl-beta-D-hexosaminidase (EC 3.2.1.52) (Fragment).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96249836; PubMed=8924513;
RA   Hall J.C., Perez F.M., Kochins J.G., Pettersen C.A., Li Y.,
RA   Tubbs C.E., LaMarche M.D.;
RT   "Quantification and localization of N-acetyl-beta-D-hexosaminidase in
RT   the adult rat testis and epididymis.";
RL   Biol. Reprod. 54:914-929(1996).
SQ   SEQUENCE   25 AA;  2722 MW;  CBC97D5B4ED97CA0 CRC64;

  Query Match             6.2%;  Score 5;  DB 11;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 4.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 SPNST 50
           |||||
Db      17 SPNST 21


RESULT 14
Q9PWR9
ID   Q9PWR9      PRELIMINARY;      PRT;    25 AA.
AC   Q9PWR9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Melanoma receptor tyrosine kinase (Fragment).
OS   Xiphophorus maculatus (Southern platyfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX   NCBI_TaxID=8083;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ONC SR-STRAIN;
RX   MEDLINE=99132631; PubMed=9931413;
RA   Schartl M., Wilde B., Hornung U.;
RT   "Triplet repeat variability in the signal peptide sequence of the Xmrk
RT   receptor tyrosine kinase gene in Xiphophorus fish.";
RL   Gene 224:17-21(1998).
DR   EMBL; U82804; AAD10123.1; -.
KW   Kinase; Receptor.
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2695 MW;  967AF74362DF4350 CRC64;

  Query Match             6.2%;  Score 5;  DB 13;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 4.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 LLVLS 7
            |||||
Db      14 LLVLS 18


RESULT 15
Q9S907
ID   Q9S907      PRELIMINARY;      PRT;    26 AA.
AC   Q9S907;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Antifungal protein (Fragment).
OS   Avena sativa (Oat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Aveneae; Avena.
OX   NCBI_TaxID=4498;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92190627; PubMed=1799695;
RA   Vigers A.J., Roberts W.K., Selitrennikoff C.P.;
RT   "A new family of plant antifungal proteins.";
RL   Mol. Plant Microbe Interact. 4:315-323(1991).
DR   InterPro; IPR001938; Thaumatin.
DR   Pfam; PF00314; thaumatin; 1.
FT   NON_TER       1      1
FT   NON_TER      26     26
SQ   SEQUENCE   26 AA;  2648 MW;  57C46A9E388E4A72 CRC64;

  Query Match             6.2%;  Score 5;  DB 10;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 4.6e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      75 PGALP 79
           |||||
Db      15 PGALP 19


Search completed: May 16, 2003, 15:35:05
Job time : 33 secs
```

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:34:30 ; Search time 29 Seconds
(without alignments)
82.181 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 51475

Minimum DB seq length: 23
Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5	6.2	23	1	US-08-086-428B-159	Sequence 159, App
2	5	6.2	23	2	US-08-468-570-159	Sequence 159, App
3	5	6.2	23	2	US-08-290-665A-263	Sequence 263, App
4	5	6.2	23	4	US-08-469-260A-400	Sequence 400, App
5	5	6.2	23	5	PCT-US95-10398-263	Sequence 263, App
6	5	6.2	24	1	US-07-928-930A-7	Sequence 7, Appli
7	5	6.2	24	1	US-08-288-568-7	Sequence 7, Appli
8	5	6.2	24	1	US-08-487-461-7	Sequence 7, Appli
9	5	6.2	24	1	US-08-432-691-7	Sequence 7, Appli
10	5	6.2	24	1	US-08-487-459-7	Sequence 7, Appli
11	5	6.2	24	4	US-09-288-143-127	Sequence 127, App
12	5	6.2	27	1	US-08-066-325-89	Sequence 89, Appl
13	5	6.2	27	4	US-08-965-762-8	Sequence 8, Appli
14	5	6.2	27	4	US-09-911-927-8	Sequence 8, Appli
15	5	6.2	27	4	US-09-911-882-8	Sequence 8, Appli
16	5	6.2	28	4	US-09-507-819-30	Sequence 30, Appl
17	5	6.2	29	1	US-08-066-325-61	Sequence 61, Appl
18	5	6.2	29	1	US-08-066-325-74	Sequence 74, Appl
19	5	6.2	29	1	US-08-066-325-75	Sequence 75, Appl
20	5	6.2	29	2	US-08-419-061A-5	Sequence 5, Appli
21	5	6.2	29	2	US-08-485-647A-5	Sequence 5, Appli
22	5	6.2	29	4	US-09-507-819-46	Sequence 46, Appl
23	5	6.2	29	4	US-09-028-274A-5	Sequence 5, Appli
24	5	6.2	30	1	US-08-066-325-60	Sequence 60, Appl
25	5	6.2	30	4	US-09-043-293-1	Sequence 1, Appli
26	5	6.2	30	4	US-09-511-023-5	Sequence 5, Appli
27	5	6.2	31	1	US-08-066-325-76	Sequence 76, Appl
28	5	6.2	31	1	US-08-066-325-77	Sequence 77, Appl
29	5	6.2	31	1	US-08-066-325-78	Sequence 78, Appl
30	5	6.2	31	4	US-09-507-819-48	Sequence 48, Appl
31	5	6.2	32	4	US-09-461-697-157	Sequence 157, App
32	5	6.2	32	4	US-09-227-357-207	Sequence 207, App
33	5	6.2	34	1	US-07-915-247A-21	Sequence 21, Appl
34	5	6.2	34	1	US-08-443-863-21	Sequence 21, Appl
35	5	6.2	34	1	US-08-448-070-21	Sequence 21, Appl
36	5	6.2	34	1	US-08-449-500-21	Sequence 21, Appl
37	5	6.2	34	1	US-08-449-317A-21	Sequence 21, Appl
38	5	6.2	34	2	US-08-477-022-21	Sequence 21, Appl
39	5	6.2	34	2	US-08-449-447-21	Sequence 21, Appl
40	5	6.2	34	2	US-08-184-328-21	Sequence 21, Appl
41	5	6.2	34	2	US-08-521-097-21	Sequence 21, Appl
42	5	6.2	35	1	US-08-449-500-84	Sequence 84, Appl
43	5	6.2	35	1	US-08-449-317A-84	Sequence 84, Appl
44	5	6.2	35	2	US-08-477-022-84	Sequence 84, Appl
45	5	6.2	35	2	US-08-449-447-84	Sequence 84, Appl

ALIGNMENTS

```
RESULT 1
US-08-086-428B-159
; Sequence 159, Application US/08086428B
; Patent No. 5514539
;  GENERAL INFORMATION:
;    APPLICANT:  BUKH, J., MILLER, R.H. AND
;    APPLICANT:  PURCELL, R.H.
;    TITLE OF INVENTION:  NUCLEOTIDE AND DEDUCED
;    TITLE OF INVENTION:  AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
;    TITLE OF INVENTION:  OF 51 ISOLATES OF HEPATITIS C AND THE USE
;    TITLE OF INVENTION:  OF REAGENTS DERIVED FROM THESE SEQUENCES IN
;    TITLE OF INVENTION:  DIAGNOSTIC METHODS AND VACCINES
;    NUMBER OF SEQUENCES:  159
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/086,428B
;      FILING DATE:  29-JUN-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  RICHARD W. BORK
;      REGISTRATION NUMBER:  36,459
;      REFERENCE/DOCKET NUMBER:  2026-4070
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  159:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
US-08-086-428B-159

  Query Match              6.2%;  Score 5;  DB 1;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 LVLSS 8
```

```
          |||||
Db      4 LVLSS 8


RESULT 2
US-08-468-570-159
; Sequence 159, Application US/08468570
; Patent No. 5871962
;  GENERAL INFORMATION:
;    APPLICANT:  BUKH, J., MILLER, R.H. AND
;    APPLICANT:  PURCELL, R.H.
;    TITLE OF INVENTION:  NUCLEOTIDE AND DEDUCED
;    TITLE OF INVENTION:  AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
;    TITLE OF INVENTION:  OF 51 ISOLATES OF HEPATITIS C AND THE USE
;    TITLE OF INVENTION:  OF REAGENTS DERIVED FROM THESE SEQUENCES IN
;    TITLE OF INVENTION:  DIAGNOSTIC METHODS AND VACCINES
;    NUMBER OF SEQUENCES:  159
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,570
;      FILING DATE:  6-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/086,428
;      FILING DATE:  29-JUN-1993
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  RICHARD W. BORK
;      REGISTRATION NUMBER:  36,459
;      REFERENCE/DOCKET NUMBER:  2026-4070US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  159:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
US-08-468-570-159

  Query Match             6.2%;  Score 5;  DB 2;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 LVLSS 8
          |||||
Db      4 LVLSS 8


RESULT 3
US-08-290-665A-263
; Sequence 263, Application US/08290665A
; Patent No. 5882852
;  GENERAL INFORMATION:
;    APPLICANT:  BUKH, J., MILLER, R.H. AND
;    APPLICANT:  PURCELL, R.H.
;    TITLE OF INVENTION:  NUCLEOTIDE AND DEDUCED
;    TITLE OF INVENTION:  AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
;    TITLE OF INVENTION:  CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
;    TITLE OF INVENTION:  AND THE USE OF REAGENTS DERIVED FROM THESE
;    TITLE OF INVENTION:  SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
;    NUMBER OF SEQUENCES:  263
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/290,665A
;      FILING DATE:  15-AUG-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  RICHARD W. BORK
;      REGISTRATION NUMBER:  36,459
;      REFERENCE/DOCKET NUMBER:  2026-4116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  263:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
US-08-290-665A-263

  Query Match             6.2%;  Score 5;  DB 2;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 LVLSS 8
          |||||
Db      4 LVLSS 8


RESULT 4
US-08-469-260A-400
; Sequence 400, Application US/08469260A
; Patent No. 6451578
;  GENERAL INFORMATION:
;    APPLICANT:  JOHN N. SIMONS
;    APPLICANT:  TAMI J. PILOT-MATIAS
;    APPLICANT:  GEORGE J. DAWSON
;    APPLICANT:  GEORGE G. SCHLAUDER
;    APPLICANT:  SURESH M. DESAI
;    APPLICANT:  THOMAS P. LEARY
;    APPLICANT:  ANTHONY SCOTT MUERHOFF
;    APPLICANT:  JAMES C. ERKER
;    APPLICANT:  SHERI L. BUIJK
;    APPLICANT:  ISA K. MUSHAHWAR
;    TITLE OF INVENTION:  NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
;    TITLE OF INVENTION:  REAGENTS AND METHODS FOR THEIR USE
;    NUMBER OF SEQUENCES:  716
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ABBOTT LABORATORIES  D377/AP6D
;      STREET:  100 ABBOTT PARK ROAD
;      CITY:  ABBOTT PARK
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,260A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/424,550
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  POREMBSKI, PRISCILLA E.
;      REGISTRATION NUMBER:  33,207
;      REFERENCE/DOCKET NUMBER:  5527.PC.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708-937-6365
;      TELEFAX:  708-938-2623
;  INFORMATION FOR SEQ ID NO:  400:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-469-260A-400

  Query Match             6.2%;  Score 5;  DB 4;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      24 GKRRP 28
           |||||
Db       8 GKRRP 12


RESULT 5
PCT-US95-10398-263
; Sequence 263, Application PC/TUS9510398
;  GENERAL INFORMATION:
;    APPLICANT:  BUKH, J., MILLER, R.H. AND
;    APPLICANT:  PURCELL, R.H.
;    TITLE OF INVENTION:  NUCLEOTIDE AND DEDUCED
;    TITLE OF INVENTION:  AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
;    TITLE OF INVENTION:  CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
;    TITLE OF INVENTION:  AND THE USE OF REAGENTS DERIVED FROM THESE
;    TITLE OF INVENTION:  SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
;    NUMBER OF SEQUENCES:  263
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/10398
;      FILING DATE:  15-AUG-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/086,428
;      FILING DATE:  29 JUNE 1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/290/665
;      FILING DATE:  15 AUGUST 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  RICHARD W. BORK
;      REGISTRATION NUMBER:  36,459
;      REFERENCE/DOCKET NUMBER:  2026-4116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  263:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
PCT-US95-10398-263

  Query Match             6.2%;  Score 5;  DB 5;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 LVLSS 8
           |||||
Db      4 LVLSS 8


RESULT 6
US-07-928-930A-7
; Sequence 7, Application US/07928930A
; Patent No. 5344822
;  GENERAL INFORMATION:
;    APPLICANT:  Levine, Daniel M.; Parker, Thomas S.;
;    APPLICANT:  Rubin, Albert L.
;    TITLE OF INVENTION:  Methods Useful in Endotoxin
;    TITLE OF INVENTION:  Prophylaxis and Therapy
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/928,930A
;      FILING DATE:  19920812
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5344822man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  ROGO 206
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-928-930A-7

  Query Match             6.2%;  Score 5;  DB 1;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 LSSLL 10
           |||||
Db      5 LSSLL 9


RESULT 7
US-08-288-568-7
; Sequence 7, Application US/08288568
; Patent No. 5506218
;  GENERAL INFORMATION:
;    APPLICANT:  Parker, Thomas S.; Levine, Daniel M.;
;    APPLICANT:  Rubin, Albert L; Gordon, Bruce R; Saal, Stuart D.
```

```
;    TITLE OF INVENTION:  Methods Useful in Endotoxin Based
;    TITLE OF INVENTION:  Prophylaxis and Therapy
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,568
;      FILING DATE:  10-AUGUST-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07453
;      FILING DATE:  9-AUGUST-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/928,930
;      FILING DATE:  12-AUGUST-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5506218man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  ROGO 211
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-288-568-7

  Query Match             6.2%;  Score 5;  DB 1;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 LSSLL 10
          |||||
Db      5 LSSLL 9


RESULT 8
US-08-487-461-7
; Sequence 7, Application US/08487461
; Patent No. 5587366
;  GENERAL INFORMATION:
;    APPLICANT:  Parker, Thomas S.; Levine, Daniel M.;
;    APPLICANT:  Rubin, Albert L; Gordon, Bruce R; Saal, Stuart D.
;    TITLE OF INVENTION:  Methods Useful in Endotoxin Based
;    TITLE OF INVENTION:  Prophylaxis and Therapy
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,461
;      FILING DATE:  June 7, 1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07453
;      FILING DATE:  9-AUGUST-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/928,930
;      FILING DATE:  12-AUGUST-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5587366man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  ROGO 211.2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-487-461-7

  Query Match             6.2%;  Score 5;  DB 1;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 LSSLL 10
          |||||
Db      5 LSSLL 9


RESULT 9
US-08-432-691-7
; Sequence 7, Application US/08432691
; Patent No. 5614507
;  GENERAL INFORMATION:
;    APPLICANT:  Parker, Thomas S.; Levine, Daniel M.;
;    APPLICANT:  Rubin, Albert L; Gordon, Bruce R; Saal, Stuart D.
;    TITLE OF INVENTION:  Methods Useful in Endotoxin Based
;    TITLE OF INVENTION:  Prophylaxis and Therapy
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/432,691
;      FILING DATE:  02-MAY-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/288,568
;      FILING DATE:  10-AUGUST-1994
;      APPLICATION NUMBER:  PCT/US93/07453
;      FILING DATE:  9-AUGUST-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/928,930
;      FILING DATE:  12-AUGUST-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5614507man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  ROGO 211
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
```

```
;       TELEFAX:  (212) 838-3884
;    INFORMATION FOR SEQ ID NO:  7:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  24 amino acids
;        TYPE:  amino acid
;        TOPOLOGY:  linear
US-08-432-691-7

  Query Match             6.2%;  Score 5;  DB 1;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 LSSLL 10
           |||||
Db       5 LSSLL 9


RESULT 10
US-08-487-459-7
; Sequence 7, Application US/08487459
; Patent No. 5674855
;  GENERAL INFORMATION:
;    APPLICANT:  Levine, Daniel M.; Parker, Thomas S.;
;    APPLICANT:  Rubin, Albert L; Gordon, Bruce R; Saal, Stuart D.
;    TITLE OF INVENTION:  Methods Useful in Endotoxin Based
;    TITLE OF INVENTION:  Prophylaxis and Therapy
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,459
;      FILING DATE:  7-JUNE-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07453
;      FILING DATE:  9-AUGUST-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/928,930
;      FILING DATE:  12-AUGUST-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5674855man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  ROGO 211.3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;   INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-487-459-7

  Query Match             6.2%;  Score 5;  DB 1;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 LSSLL 10
           |||||
Db       5 LSSLL 9


RESULT 11
US-09-288-143-127
; Sequence 127, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
;  APPLICANT: Brewer et al.
;  TITLE OF INVENTION: 53 Human Secreted Proteins
;  FILE REFERENCE: PZ018P1
;  CURRENT APPLICATION NUMBER: US/09/288,143
;  CURRENT FILING DATE: 1999-04-08
;  EARLIER APPLICATION NUMBER: PCT/US98/21142
;  EARLIER FILING DATE: 1998-10-08
;  EARLIER APPLICATION NUMBER: 60/061,463
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,529
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/071,498
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,527
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,536
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,532
;  EARLIER FILING DATE: 1997-10-09
;  NUMBER OF SEQ ID NOS: 219
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 127
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-288-143-127

  Query Match             6.2%;  Score 5;  DB 4;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 LVLSS 8
           |||||
Db       9 LVLSS 13


RESULT 12
US-08-066-325-89
; Sequence 89, Application US/08066325
; Patent No. 5667967
;  GENERAL INFORMATION:
;    APPLICANT:  Steinman, Lawrence
;    APPLICANT:  Oksenberg, Jorge
;    APPLICANT:  Bernard, Claude
;    TITLE OF INVENTION:  T-CELL RECEPTOR VARIABLE TRANSCRIPTS AS DISEASE RELATED MARKE
;    NUMBER OF SEQUENCES:  157
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/066,325
;      FILING DATE:  21-MAY-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5667967tenburg Ph.D., Carol
;      REGISTRATION NUMBER:  39,317
;      REFERENCE/DOCKET NUMBER:  690068.408C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
```

```
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-066-325-89

  Query Match             6.2%;  Score 5;  DB 1;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 RLLVL 6
           |||||
Db      18 RLLVL 22


RESULT 13
US-08-965-762-8
; Sequence 8, Application US/08965762
; Patent No. 6280963
; GENERAL INFORMATION:
;  APPLICANT: Koltin, Yigal
;  APPLICANT: Gavrias, Victoria
;  TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
;  FILE REFERENCE: 07334/062001
;  CURRENT APPLICATION NUMBER: US/08/965,762
;  CURRENT FILING DATE: 1997-11-07
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 27
;   TYPE: PRT
;   ORGANISM: Aspergillus nidulans
US-08-965-762-8

  Query Match             6.2%;  Score 5;  DB 4;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 CILLL 15
           |||||
Db       4 CILLL 8


RESULT 14
US-09-911-927-8
; Sequence 8, Application US/09911927
; Patent No. 6461826
; GENERAL INFORMATION:
;  APPLICANT: Koltin, Yigal
;  APPLICANT:  Gavrias, Victoria
;  TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
;  FILE REFERENCE: 06286-062003
;  CURRENT APPLICATION NUMBER: US/09/911,927
;  CURRENT FILING DATE:  2001-07-23
;  PRIOR APPLICATION NUMBER: US 08/965,762
;  PRIOR FILING DATE: 1997-11-07
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 27
;   TYPE: PRT
;   ORGANISM: Aspergillus nidulans
US-09-911-927-8

  Query Match             6.2%;  Score 5;  DB 4;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 CILLL 15
           |||||
Db       4 CILLL 8


RESULT 15
US-09-911-882-8
; Sequence 8, Application US/09911882
; Patent No. 6465198
; GENERAL INFORMATION:
;  APPLICANT: Koltin, Yigal
;  APPLICANT:  Gavrias, Victoria
;  TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
;  FILE REFERENCE: 06286-062004
;  CURRENT APPLICATION NUMBER: US/09/911,882
;  CURRENT FILING DATE:  2001-07-23
;  PRIOR APPLICATION NUMBER: US 08/965,762
;  PRIOR FILING DATE: 1997-11-07
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 27
;   TYPE: PRT
;   ORGANISM: Aspergillus nidulans
US-09-911-882-8

  Query Match             6.2%;  Score 5;  DB 4;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 CILLL 15
           |||||
Db       4 CILLL 8


Search completed: May 16, 2003, 15:36:11
Job time : 38 secs
```

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2003, 03:26:33 ; Search time 2182 Seconds
(without alignments)
1080.352 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 442
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09724000/runat_16052003_145459_17144/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09724000_@CGN_1_1_2496_@runat_16052003_145459_17144 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	442	100.0	382	6	AX340424	AX340424 Sequence
2	442	100.0	485	6	AX351259	AX351259 Sequence
3	442	100.0	804	6	AX464016	AX464016 Sequence
4	442	100.0	806	6	AX342218	AX342218 Sequence
5	442	100.0	2063	9	AK025416	AK025416 Homo sapi
c 6	400	90.5	797	6	AX027773	AX027773 Sequence
c 7	400	90.5	801	6	AX027767	AX027767 Sequence
8	228.5	51.7	744	6	AX342215	AX342215 Sequence
9	183	41.4	744	10	AF152002	AF152002 Rattus no
10	178	40.3	4159	6	AX342222	AX342222 Sequence
11	178	40.3	198481	9	AC022389	AC022389 Homo sapi
12	160.5	36.3	742	10	S74257	S74257 2c9 gene {c
c 13	111.5	25.2	137164	2	AC128725	AC128725 Rattus no
14	106.5	24.1	225929	2	AC109260	AC109260 Mus muscu
c 15	92	20.8	207374	9	CNS07EFU	AL583722 Human chr
c 16	88.5	20.0	173235	2	AC106276	AC106276 Rattus no
17	87.5	19.8	118429	2	AL158017	AL158017 Homo sapi
c 18	87.5	19.8	198906	9	AL356275	AL356275 Human DNA
19	86.5	19.6	2167	6	A37236	A37236 Sequence 5
20	86.5	19.6	2168	4	BOVPKG	M13976 Bovine gamm
21	86.5	19.6	108071	2	AC123218	AC123218 Rattus no
22	86.5	19.6	161044	2	AC119102	AC119102 Rattus no
23	86.5	19.6	176092	2	AC073252	AC073252 Homo sapi
c 24	86	19.5	170000	2	AC004394	AC004394 Homo sapi
c 25	86	19.5	170000	2	AC004579	AC004579 Homo sapi
26	86	19.5	190000	2	AC004580	AC004580 Homo sapi
c 27	86	19.5	210515	9	AC097382	AC097382 Homo sapi
28	85.5	19.3	10647	1	AE002031	AE002031 Deinococc
29	85.5	19.3	110000	2	LMFLCHR36_10	Continuation (11 o
c 30	85.5	19.3	161112	2	AC115876	AC115876 Mus muscu
31	85	19.2	853	9	BC002624	BC002624 Homo sapi
32	85	19.2	876	9	BC001181	BC001181 Homo sapi
c 33	85	19.2	1688	6	AR035537	AR035537 Sequence
c 34	85	19.2	1688	6	E11457	E11457 cDNA encodi
c 35	85	19.2	1934	9	BC007790	BC007790 Homo sapi
c 36	85	19.2	1979	6	AR035536	AR035536 Sequence
c 37	85	19.2	1979	6	E11456	E11456 cDNA encodi
c 38	85	19.2	2010	9	HSU07132	U07132 Human stero
c 39	85	19.2	2030	6	I36667	I36667 Sequence 1
c 40	85	19.2	2030	6	I70211	I70211 Sequence 1
c 41	85	19.2	2384	9	AK091535	AK091535 Homo sapi
c 42	85	19.2	157375	2	AC122095	AC122095 Rattus no
c 43	85	19.2	195983	2	AC099699	AC099699 Mus muscu
c 44	84.5	19.1	2140	10	AF163665	AF163665 Mus muscu
c 45	84.5	19.1	2398	10	AF282322	AF282322 Mus muscu

ALIGNMENTS

RESULT 1

```
AX340424/c
LOCUS       AX340424                382 bp    DNA     linear   PAT 10-JAN-2002
DEFINITION  Sequence 671 from Patent WO0196388.
ACCESSION   AX340424
VERSION     AX340424.1  GI:18136406
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Harlocker,S.L. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196388-A 671 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       76 a     94 c    122 g     86 t      4 others
ORIGIN

Alignment Scores:
Pred. No.:              1.22e-33        Length:         382
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-724-000A-5 (1-81) x AX340424 (1-382)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 242

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 182

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 122

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 62

Qy     81 Val 81
          |||
Db     61 GTG 59

RESULT 2
AX351259
LOCUS       AX351259                485 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 6 from Patent WO0196390.
ACCESSION   AX351259
VERSION     AX351259.1  GI:18616606
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196390-A 6 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      109 a    155 c    124 g     97 t
ORIGIN

Alignment Scores:
Pred. No.:              1.58e-33        Length:         485
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-724-000A-5 (1-81) x AX351259 (1-485)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 138

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 198

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 258

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 318

Qy     81 Val 81
          |||
Db    319 GTG 321

RESULT 3
AX464016
LOCUS       AX464016                804 bp    DNA     linear   PAT 16-JUL-2002
DEFINITION  Sequence 149 from Patent WO0140466.
ACCESSION   AX464016
VERSION     AX464016.1  GI:21899025
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,
            Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
            Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
            Wood,W.L. and Zhang,Z.
  TITLE     Secreted and transmembrane polypeptides and nucleic acids encoding
            same
  JOURNAL   Patent: WO 0140466-A 149 07-JUN-2001;
            Genentech Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .804
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      199 a    263 c    176 g    166 t
ORIGIN

Alignment Scores:
Pred. No.:              2.74e-33        Length:         804
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-724-000A-5 (1-81) x AX464016 (1-804)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     42 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 101

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    102 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 161

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 281

Qy     81 Val 81
          |||
Db    282 GTG 284

RESULT 4
AX342218
LOCUS       AX342218                806 bp    DNA     linear   PAT 12-JAN-2002
DEFINITION  Sequence 4 from Patent WO0198497.
ACCESSION   AX342218
VERSION     AX342218.1  GI:18151763
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Polverino,A.J. and Luethy,R.
  TITLE     Secreted epithelial colon stromal-1 polypeptides, nucleic acids
            encoding the same and uses thereof
  JOURNAL   Patent: WO 0198497-A 4 27-DEC-2001;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .806
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             29. .274
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD20706.1"
                     /db_xref="GI:18151764"
                     /translation="MRLLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRV
                     PSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPGALPQV"
     sig_peptide     29. .100
BASE COUNT      207 a    257 c    179 g    163 t
ORIGIN

Alignment Scores:
Pred. No.:              2.75e-33        Length:         806
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-724-000A-5 (1-81) x AX342218 (1-806)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     29 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 88

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     89 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 148

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    149 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 208

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    209 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 268

Qy     81 Val 81
          |||
Db    269 GTG 271

RESULT 5
AK025416
LOCUS       AK025416                2063 bp    mRNA    linear   PRI 29-SEP-2000
DEFINITION  Homo sapiens cDNA: FLJ21763 fis, clone COLF6967.
ACCESSION   AK025416
VERSION     AK025416.1  GI:10437924
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens colon mucosa cDNA to mRNA, clone_lib:ColF
            clone:COLF6967.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
            Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2063)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .2063
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="COLF6967"
                     /tissue_type="colon mucosa"
                     /clone_lib="ColF"
                     /note="cloning vector pME18SFL3"
BASE COUNT      536 a    479 c    476 g    572 t
ORIGIN

Alignment Scores:
Pred. No.:              7.59e-33        Length:         2063
Score:                  442.00          Matches:        81
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-724-000A-5 (1-81) x AK025416 (1-2063)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    115 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 174

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    175 TCCACAGAAGGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGC 234

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    235 TGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 294
```

```
Qy      61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     295 AAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGCACTCCCACAG 354

Qy      81 Val 81
           |||
Db     355 GTG 357

RESULT 6
AX027773/c
LOCUS       AX027773                797 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 7 from Patent WO0043509.
ACCESSION   AX027773
VERSION     AX027773.1  GI:10188625
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 797)
  AUTHORS   Vinals-Bassols,C.
  TITLE     Novel compounds
  JOURNAL   Patent: WO 0043509-A 7 27-JUL-2000;
            SMITHKLINE BEECHAM BIOLOG (BE) ; VINALS BASSOLS CARLOTA (BE)
FEATURES             Location/Qualifiers
     source          1. .797
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      164 a    180 c    254 g    199 t
ORIGIN

Alignment Scores:
Pred. No.:              2.71e-29        Length:         797
Score:                  400.00          Matches:        78
Percent Similarity:     96.34%          Conservative:   1
Best Local Similarity:  95.12%          Mismatches:     2
Query Match:            90.50%          Indels:         2
DB:                     6               Gaps:           0

US-09-724-000A-5 (1-81) x AX027773 (1-797)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     765 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 706

Qy      21 SerThrGluGlyLysArgArgProAlaLys-AlaTrpSerGlyArgArgThrArgLeuCy 40
           ||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db     705 TCCACAGAGGGGAAGAGGCGTCCTGCCAACGGCCTGGTCAGGCAGGAGAACCAGGCTCTG 646

Qy      40 sCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCy 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     645 CTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTG 586

Qy      60 sLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGl 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||   :::|||||
Db     585 TAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGA-GTCCCACA 527

Qy      80 nVal 81
           ||||
Db     526 GGTG 523

RESULT 7
AX027767/c
LOCUS       AX027767                801 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 1 from Patent WO0043509.
ACCESSION   AX027767
VERSION     AX027767.1  GI:10188619
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   Vinals-Bassols,C.
  TITLE     Novel compounds
  JOURNAL   Patent: WO 0043509-A 1 27-JUL-2000;
            SMITHKLINE BEECHAM BIOLOG (BE) ; VINALS BASSOLS CARLOTA (BE)
FEATURES             Location/Qualifiers
     source          1. .801
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      165 a    181 c    255 g    200 t
ORIGIN

Alignment Scores:
Pred. No.:              2.72e-29        Length:         801
Score:                  400.00          Matches:        78
Percent Similarity:     96.34%          Conservative:   1
Best Local Similarity:  95.12%          Mismatches:     2
Query Match:            90.50%          Indels:         2
DB:                     6               Gaps:           0

US-09-724-000A-5 (1-81) x AX027767 (1-801)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 ATGAGGCTTCTAGTCCTTTCCAGCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTC 710

Qy      21 SerThrGluGlyLysArgArgProAlaLys-AlaTrpSerGlyArgArgThrArgLeuCy 40
           ||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db     709 TCCACAGAGGGGAAGAGGCGTCCTGCCAACGGCCTGGTCAGGCAGGAGAACCAGGCTCTG 650

Qy      40 sCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCy 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 CTGCCACCGAGTCCCTAGCCCCAACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTG 590

Qy      60 sLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGl 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||   :::|||||
Db     589 TAAACCATGCAAGCTTGAGCCAGAGCCCCGCCTTTGGGTGGTGCCTGGGGA-GTCCCACA 531

Qy      80 nVal 81
           ||||
Db     530 GGTG 527

RESULT 8
AX342215
LOCUS       AX342215                744 bp    DNA     linear   PAT 12-JAN-2002
DEFINITION  Sequence 1 from Patent WO0198497.
ACCESSION   AX342215
VERSION     AX342215.1  GI:18151761
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Polverino,A.J. and Luethy,R.
  TITLE     Secreted epithelial colon stromal-1 polypeptides, nucleic acids
            encoding the same and uses thereof
  JOURNAL   Patent: WO 0198497-A 1 27-DEC-2001;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             38. .274
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD20705.1"
                     /db_xref="GI:18151762"
                     /translation="MRLLALSGLLCMLLLCFCIFSSEGRRHPAKSLKLRRCCHLSPRS
                     KLTTWKGNHTRPCRLCRNKLPVKSWVVPGALPQI"
     sig_peptide     38. .109
BASE COUNT      209 a    222 c    155 g    158 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              5.39e-13        Length:         744
Score:                  228.50          Matches:        45
Percent Similarity:     67.90%          Conservative:   10
Best Local Similarity:  55.56%          Mismatches:     23
Query Match:            51.70%          Indels:         3
DB:                     6               Gaps:           1

US-09-724-000A-5 (1-81) x AX342215 (1-744)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||   ||||||  |||||||||:::|||||||||||||||   ||||||
Db     38 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 97

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          |||:::||||||:::|||  |||||||||:::       ||||||          |||
Db     98 TCCTCAGAAGGGAGAAGACATCCTGCCAAGTCCTTGAAACTCAGGCGC---------TGC 148

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||      |||             |||      ||||||:::|||   |||   |||
Db    149 TGTCACCTATCTCCTAGATCCAAGCTGACAACCTGGAAAGGAAACCACACAAGGCCCTGC 208

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          :::   |||:::   :::         :::   ||||||||||||||||||||||||||
Db    209 AGACTCTGCAGAAACAAGCTACCAGTCAAGTCATGGGTGGTGCCTGGGGCTCTCCCACAG 268

Qy     81 Val 81
          :::
Db    269 ATA 271

RESULT 9
AF152002
LOCUS       AF152002                744 bp    mRNA    linear   ROD 29-JUN-1999
DEFINITION  Rattus norvegicus unknown mRNA sequence.
ACCESSION   AF152002
VERSION     AF152002.1  GI:5257466
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Rossman,T.G. and Li,P.
  TITLE     2C9-like sequence expressed in lead-resistant rat glioma cells
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 744)
  AUTHORS   Rossman,T.G. and Li,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-1999) Environmental Medicine, NYU School of
            Medicine, 57 Old Forge, Tuxedo, NY 10987, USA
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /cell_line="PbR1"
                     /tissue_type="glioma"
                     /note="lead resistant cells derived from C6 cell line;
                     Dolzhanskaya et al., 1998, Biological Trace Element Res.
                     65:31-43"
     misc_feature    1. .744
                     /note="sequence expressed in PbR1 cells; not expressed in
                     C6 cells; sequence resembles 2C9, found in cells
                     overexpressing fos; PbR1 cells do not overexpress fos"
BASE COUNT      208 a    216 c    146 g    174 t
ORIGIN

Alignment Scores:
Pred. No.:              1.16e-08        Length:         744
Score:                  183.00          Matches:        38
Percent Similarity:     60.49%          Conservative:   11
Best Local Similarity:  46.91%          Mismatches:     28
Query Match:            41.40%          Indels:         4
DB:                     10              Gaps:           2

US-09-724-000A-5 (1-81) x AF152002 (1-744)

Qy      1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
          ||||||||||||   ||||||  |||      :::|||   ||||||      :::
Db     53 ATGCGACTTCTCACCCTCTCCGGTTTGTTCTTCATGCTGTTCCTCTGTCTCTGCGTTCTC 112

Qy     21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
          |||:::||||||:::::||||||||||                :::   |||
Db    113 TCCTCAGAAGGGAGAAAGCGTCCTGCCAAGTTC---------CCGAAACTCAGGCCCCGC 163

Qy     41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
          ||||||      |||                  ||||||:::|||   |||   |||
Db    164 TGTCATCTATCTCCTAGATCCAAACCAATAACCTGGAAAGGAAACCACACAAGACCCTGC 223

Qy     61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
          :::|||||:::   :::   |||         ||||||||||||||||||||||||||
Db    224 AGACCATGCAGA---AAGCTAGAATCCAATTCATGGGTGGTGCCTGGGGCTCTCCCACAG 280

Qy     81 Val 81
          :::
Db    281 ATA 283

RESULT 10
AX342222
LOCUS       AX342222                4159 bp    DNA     linear   PAT 12-JAN-2002
DEFINITION  Sequence 8 from Patent WO0198497.
ACCESSION   AX342222
VERSION     AX342222.1  GI:18151765
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Polverino,A.J. and Luethy,R.
  TITLE     Secreted epithelial colon stromal-1 polypeptides, nucleic acids
            encoding the same and uses thereof
  JOURNAL   Patent: WO 0198497-A 8 27-DEC-2001;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4159
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     exon            1. .69
     exon            2627. .2725
     exon            4079. .4159
BASE COUNT     1009 a   1037 c    983 g   1110 t     20 others
ORIGIN

Alignment Scores:
Pred. No.:              2.23e-07        Length:         4159
Score:                  178.00          Matches:        37
Percent Similarity:     62.12%          Conservative:   4
Best Local Similarity:  56.06%          Mismatches:     13
Query Match:            40.27%          Indels:         13
DB:                     6               Gaps:           1

US-09-724-000A-5 (1-81) x AX342222 (1-4159)

Qy     24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArg 43
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2635 GGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGA 2694

Qy     44 ValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCysLysProCys 63
          ||||||||||||||||||||||||||||||||   :::      |||   :::   |||
Db   2695 GTCCCTAGCCCCAACTCAACAAACCTGAAAGGTAAGTACCCCCACCTCGTCCAGAC-TGT 2753

Qy     64 LysLeuGluPro----------------------------------GluProArgLeu 71
          |||                                            |||
Db   2754 GGGGCAGAAGTTCTACAGTGGCCATGGGACCAGCCACACACACTGATCAGCCCCCACCCA 2813
```

```
Qy      72 TrpValValProGlyAla 77
           |||:::      |||:::
Db    2814 TGGCTGGCATCAGGCTCT 2831

RESULT 11
AC022389
LOCUS       AC022389              198481 bp    DNA     linear   PRI 04-JUL-2002
DEFINITION  Homo sapiens chromosome 10 clone RP11-124L5, complete sequence.
ACCESSION   AC022389
VERSION     AC022389.9  GI:21321764
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2000) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   3  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   4  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUN-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   5  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   6  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   7  (bases 1 to 198481)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Jun 4, 2002 this sequence version replaced gi:19071578.
FEATURES             Location/Qualifiers
     source          1. .198481
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-124L5"
                     /clone_lib="RPCI-11"
BASE COUNT    55275 a  44232 c  45110 g  53864 t
ORIGIN

Alignment Scores:
Pred. No.:              1.46e-05        Length:         198481
Score:                  178.00          Matches:        37
Percent Similarity:     62.12%          Conservative:   4
Best Local Similarity:  56.06%          Mismatches:     13
Query Match:            40.27%          Indels:         13
DB:                     9               Gaps:           1

US-09-724-000A-5 (1-81) x AC022389 (1-198481)

Qy      24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArg 43
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34893 GGGAAGAGGCGTCCTGCCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGA 34952

Qy      44 ValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCysLysProCys 63
           |||||||||||||||||||||||||||||||||   :::      |||   :::   |||
Db   34953 GTCCCTAGCCCCAACTCAACAAACCTGAAAGGTAAGTACCCCCACCTCGTCCAGAC-TGT 35011

Qy      64 LysLeuGluPro-----------------------------------GluProArgLeu 71
           |||                                            |||
Db   35012 GGGGCAGAAGTTCTACAGTGGCCATGGGACCAGCCACACACACTGATCAGCCCCCACCCA 35071

Qy      72 TrpValValProGlyAla 77
           |||:::      |||:::
Db   35072 TGGCTGGCATCAGGCTCT 35089

RESULT 12
S74257
LOCUS       S74257                 742 bp    mRNA    linear   ROD 16-MAR-1995
DEFINITION  2c9 gene {clone 2C9 incert} [rats, rat-1 fibroblast 208F,
            Fos-transformed, mRNA, 742 nt].
ACCESSION   S74257
VERSION     S74257.1  GI:710462
KEYWORDS    .
SOURCE      Rattus sp. Fos-transformed rat-1 fibroblast 208F.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 742)
  AUTHORS   Hennigan,R.F., Hawker,K.L. and Ozanne,B.W.
  TITLE     Fos-transformation activates genes associated with invasion
  JOURNAL   Oncogene 9 (12), 3591-3600 (1994)
  MEDLINE   95060817
   PUBMED   7970719
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 157814] from the original journal article.
            This sequence comes from Fig. 5.
FEATURES             Location/Qualifiers
     source          1. .742
                     /organism="Rattus sp."
                     /db_xref="taxon:10118"
     gene            1. .742
                     /gene="2c9 gene"
BASE COUNT      208 a    213 c    146 g    175 t
ORIGIN

Alignment Scores:
Pred. No.:              1.61e-06        Length:         742
Score:                  160.50          Matches:        40
Percent Similarity:     63.41%          Conservative:   12
Best Local Similarity:  48.78%          Mismatches:     25
Query Match:            36.31%          Indels:         7
DB:                     10              Gaps:           2

US-09-724-000A-5 (1-81) x S74257 (1-742)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||   ||||||   |||       :::|||   ||||||      :::
Db     53 ATGCGACTTCTCACCCTCTCCGGTTTGTTCTTCATGCTGTTCCTCTGTCTCTGCGTTCTC 112

Qy      21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
           |||:::||||||::::::||||||||||||                  ||||:::   |||
Db    113 TCCTCAGAAGGGAGAAAGCGTCCTGCCAAGTT----------CCCGAAACTCAGGCCTGC 162

Qy      41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
           |||||||     |||       :::   |||   ||||||:::|||   |||   |||
Db    163 TGTCATCTATCTCCTAGATC-CAAACCAATAACTGGAAAGGAAACCACACAAGACCCTGC 221

Qy      61 LysProCysLysLeuGluProGluProArgLeuTrp-ValValProGlyAlaLeuProGl 80
           :::||||||:::   :::   |||         ||| ||||||||||||||||||||||
Db    222 AGACCATGCAGA---AAGCTAGAATCCAATTCATGGGGTGGTGCCTGGGGCTCTCCCACA 278

Qy      80 nVal 81
```

```
          |:::
Db    279 GATA 282

RESULT 13
AC128725/c
LOCUS       AC128725               137164 bp    DNA     linear   HTG 23-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-300I4, *** SEQUENCING IN PROGRESS
            ***, 54 unordered pieces.
ACCESSION   AC128725
VERSION     AC128725.1  GI:21930188
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 137164)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 137164)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GSIT
                Center clone name: CH230-300I4
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 81651 bases at least Q40
                Consensus quality: 85838 bases at least Q30
                Consensus quality: 89650 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 54 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1055: contig of 1055 bp in length
            *     1056     1155: gap of unknown length
            *     1156     2253: contig of 1098 bp in length
            *     2254     2353: gap of unknown length
            *     2354     3835: contig of 1482 bp in length
            *     3836     3935: gap of unknown length
            *     3936     5004: contig of 1069 bp in length
            *     5005     5104: gap of unknown length
            *     5105     7027: contig of 1923 bp in length
            *     7028     7127: gap of unknown length
            *     7128     8380: contig of 1253 bp in length
            *     8381     8480: gap of unknown length
            *     8481     9517: contig of 1037 bp in length
            *     9518     9617: gap of unknown length
            *     9618    11124: contig of 1507 bp in length
            *    11125    11224: gap of unknown length
            *    11225    12224: contig of 1000 bp in length
            *    12225    12324: gap of unknown length
            *    12325    13444: contig of 1120 bp in length
            *    13445    13544: gap of unknown length
            *    13545    14548: contig of 1004 bp in length
            *    14549    14648: gap of unknown length
            *    14649    16065: contig of 1417 bp in length
            *    16066    16165: gap of unknown length
            *    16166    17177: contig of 1012 bp in length
            *    17178    17277: gap of unknown length
            *    17278    18704: contig of 1427 bp in length
            *    18705    18804: gap of unknown length
            *    18805    19895: contig of 1091 bp in length
            *    19896    19995: gap of unknown length
            *    19996    21433: contig of 1438 bp in length
            *    21434    21533: gap of unknown length
            *    21534    22767: contig of 1234 bp in length
            *    22768    22867: gap of unknown length
            *    22868    24068: contig of 1201 bp in length
            *    24069    24168: gap of unknown length
            *    24169    25784: contig of 1616 bp in length
            *    25785    25884: gap of unknown length
            *    25885    27104: contig of 1220 bp in length
            *    27105    27204: gap of unknown length
            *    27205    28209: contig of 1005 bp in length
            *    28210    28309: gap of unknown length
            *    28310    30046: contig of 1737 bp in length
            *    30047    30146: gap of unknown length
            *    30147    31379: contig of 1233 bp in length
            *    31380    31479: gap of unknown length
            *    31480    33395: contig of 1916 bp in length
            *    33396    33495: gap of unknown length
            *    33496    35282: contig of 1787 bp in length
            *    35283    35382: gap of unknown length
            *    35383    38059: contig of 2677 bp in length
            *    38060    38159: gap of unknown length
            *    38160    40027: contig of 1868 bp in length
            *    40028    40127: gap of unknown length
            *    40128    42034: contig of 1907 bp in length
```

```
          *    42035    42134: gap of unknown length
          *    42135    43770: contig of 1636 bp in length
          *    43771    43870: gap of unknown length
          *    43871    45555: contig of 1685 bp in length
          *    45556    45655: gap of unknown length
          *    45656    46699: contig of 1044 bp in length
          *    46700    46799: gap of unknown length
          *    46800    49345: contig of 2546 bp in length
          *    49346    49445: gap of unknown length
          *    49446    50569: contig of 1124 bp in length
          *    50570    50669: gap of unknown length
          *    50670    53663: contig of 2994 bp in length
          *    53664    53763: gap of unknown length
          *    53764    55762: contig of 1999 bp in length
          *    55763    55862: gap of unknown length
          *    55863    57537: contig of 1675 bp in length
          *    57538    57637: gap of unknown length
          *    57638    59870: contig of 2233 bp in length
          *    59871    59970: gap of unknown length
          *    59971    61894: contig of 1924 bp in length
          *    61895    61994: gap of unknown length
          *    61995    64368: contig of 2374 bp in length
          *    64369    64468: gap of unknown length
          *    64469    67294: contig of 2826 bp in length
          *    67295    67394: gap of unknown length
          *    67395    70448: contig of 3054 bp in length
          *    70449    70548: gap of unknown length
          *    70549    72904: contig of 2356 bp in length
          *    72905    73004: gap of unknown length
          *    73005    77021: contig of 4017 bp in length
          *    77022    77121: gap of unknown length
          *    77122    78965: contig of 1844 bp in length
          *    78966    79065: gap of unknown length
          *    79066    82124: contig of 3059 bp in length
          *    82125    82224: gap of unknown length
          *    82225    86402: contig of 4178 bp in length
          *    86403    86502: gap of unknown length
          *    86503    92107: contig of 5605 bp in length
          *    92108    92207: gap of unknown length
          *    92208    96517: contig of 4310 bp in length
          *    96518    96617: gap of unknown length
          *    96618   101075: contig of 4458 bp in length
          *   101076   101175: gap of unknown length
          *   101176   108002: contig of 6827 bp in length
          *   108003   108102: gap of unknown length
          *   108103   115768: contig of 7666 bp in length
          *   115769   115868: gap of unknown length
          *   115869   122147: contig of 6279 bp in length
          *   122148   122247: gap of unknown length
          *   122248   129522: contig of 7275 bp in length
          *   129523   129622: gap of unknown length
          *   129623   137164: contig of 7542 bp in length.
FEATURES             Location/Qualifiers
     source          1. .137164
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-300I4"
BASE COUNT    37001 a  29055 c  29355 g  34732 t   7021 others
ORIGIN

Alignment Scores:
Pred. No.:              21.1          Length:        137164
Score:                  111.50        Matches:       27
Percent Similarity:     42.50%        Conservative:  7
Best Local Similarity:  33.75%        Mismatches:    15
Query Match:            25.23%        Indels:        31
DB:                     2             Gaps:          3

US-09-724-000A-5 (1-81) x AC128725 (1-137164)

Qy      2 ArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePheSer 21
          |||  |||     |||     |||   :::   ||||||||
Db  42764 CGTTTTCTACCCTCTTCACTTTGCCTTCCCGTGCCTCTTCTTTGT-------------- 42720

Qy     22 ThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCys 41
                                                    :::||||||
Db  42719 ----------------------------------------------------GTCTGCTGT 42711

Qy     42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCysLys 61
          |||      |||                        |||:::|||   |||   |||:::
Db  42710 CACTCATGTCCA----------------------GGAAACCACACAAGACCCTGCAGA 42675

Qy     62 ProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGlnVal 81
          ||||||:::   :::   |||         |||||||||||||||||||||||||||:::
Db  42674 CCATGCAGA---AAGCTAGAATCCAATTCATGGGTGGTGCCTGGGGCTCTCCCACAGATA 42618

RESULT 14
AC109260
LOCUS       AC109260               225929 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Mus musculus clone RP23-368G21, WORKING DRAFT SEQUENCE, 10 ordered
            pieces.
ACCESSION   AC109260
VERSION     AC109260.4  GI:22474867
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 225929)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP23-368G21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 225929)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 225929)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
```

```
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Aug 24, 2002 this sequence version replaced gi:20336102.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L20912
                Center clone name: 368_G_21
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 223420 bases at least Q40
                Consensus quality: 224295 bases at least Q30
                Consensus quality: 224703 bases at least Q20
                Insert size: 210000; agarose-fp
                Insert size: 225029; sum-of-contigs
                Quality coverage: 13.3 in Q20 bases; agarose-fp
                Quality coverage: 12.4 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1     2073: contig of 2073 bp in length
            *     2074 2173: gap of     100 bp
            *     2174     3695: contig of 1522 bp in length
            *     3696 3795: gap of      100 bp
            *     3796     7937: contig of 4142 bp in length
            *     7938 8037: gap of      100 bp
            *     8038    11164: contig of 3127 bp in length
            *    11165 11264: gap of      100 bp
            *    11265    27129: contig of 15865 bp in length
            *    27130 27229: gap of      100 bp
            *    27230    42139: contig of 14910 bp in length
            *    42140 42239: gap of      100 bp
            *    42240    59229: contig of 16990 bp in length
            *    59230 59329: gap of      100 bp
            *    59330    78798: contig of 19469 bp in length
            *    78799 78898: gap of      100 bp
            *    78899   144433: contig of 65535 bp in length
            *   144434 144533: gap of      100 bp
            *   144534   225929: contig of 81396 bp in length.
FEATURES             Location/Qualifiers
     source          1. .225929
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="RP23-368G21"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature    1. .2073
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    2174. .3695
                     /note="assembly_fragment"
     misc_feature    3796. .7937
                     /note="assembly_fragment"
     misc_feature    8038. .11164
                     /note="assembly_fragment"
     misc_feature    11265. .27129
                     /note="assembly_fragment"
     misc_feature    27230. .42139
                     /note="assembly_fragment"
     misc_feature    42240. .59229
                     /note="assembly_fragment"
     misc_feature    59330. .78798
                     /note="assembly_fragment"
     misc_feature    78899. .144433
                     /note="assembly_fragment"
     misc_feature    144534. .225929
                     /note="assembly_fragment"
BASE COUNT    67274 a  46621 c  46543 g  64591 t    900 others
ORIGIN

Alignment Scores:
Pred. No.:              108          Length:          225929
Score:                  106.50       Matches:         24
Percent Similarity:     60.87%       Conservative:    4
Best Local Similarity:  52.17%       Mismatches:      15
Query Match:            24.10%       Indels:          3
DB:                     2            Gaps:            1

US-09-724-000A-5 (1-81) x AC109260 (1-225929)

Qy       1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
           ||||||||||||   ||||||   |||||||||:::|||||||||||||||   ||||||
Db   24196 ATGAGACTTCTAGCCCTTTCCGGTCTGCTCTGCATGCTGCTCCTCTGTTTCTGCATTTTC 24255

Qy      21 SerThrGluGlyLysArgArg---------ProAlaLysAlaTrpSerGlyArgArgThr 37
           |||:::||||||                      :::      |||   |||
Db   24256 TCCTCAGAAGGTAACCTGCCCCTGGGCATTGACTCCGTGGTCTGGGAAGGTTTTGAACAG 24315

Qy      38 ArgLeuCysCysHisArg 43
           :::|||     ||||||
Db   24316 AAACTAGTGGGGCATAGG 24333

RESULT 15
CNS07EFU/c
LOCUS       CNS07EFU              207374 bp    DNA     linear   PRI 06-JUL-2002
DEFINITION  Human chromosome 14 DNA sequence BAC R-982M15 of library RPCI-11
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL583722
VERSION     AL583722.5  GI:21712965
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 207374)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 207374)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2002) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Jul 8, 2002 this sequence version replaced gi:15526346.
            -------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            --------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            -------------- Finishing boundaries
            FINISHED SEGMENT STARTS AT BASE 34250
            FINISHED SEGMENT ENDS AT BASE 211616
            --------------
```

```
          ---------- Summary Statistics
              Assembly program: Phrap; version 2.0
          ----------
          ----------
          Overall quality chart :
          Range    :   bases
          0        :
          1  -  9  :       8
          10 - 19  :      52
          20 - 29  :     194
          30 - 39  :     888
          40 - 49  :    5766
          50 - 59  :   10291
          60 - 69  :   11885
          70 - 79  :   24619
          80 - 89  :   59705
          90 - 99  :   93966
          ----------
          Percentage of bases with a quality value >= 40 : 99 %.
FEATURES             Location/Qualifiers
     source          1. .207374
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="R-982M15"
                     /clone_lib="RPCI-11"
     repeat_region   48734. .50176
                     /evidence=not_experimental
                     /insertion_seq="IS150"
     STS             123536. .123658
                     /note="matching EMBL:AA431297
                     RHdb:RH98436
                     dbSTS:STS68194
                     Identified using the e-PCR software (G. Schuler)"
     STS             127693. .127862
                     /note="matching EMBL:AA621795
                     RHdb:RH102965
                     dbSTS:STS70502
                     Identified using the e-PCR software (G. Schuler)"
     STS             155522. .155646
                     /note="matching EMBL:G19662
                     RHdb:RH10471
                     dbSTS:STS13897
                     Identified using the e-PCR software (G. Schuler)"
     STS             155534. .155683
                     /note="matching EMBL:H82735
                     RHdb:RH53711
                     dbSTS:STS8268
                     Identified using the e-PCR software (G. Schuler)"
     STS             177768. .177923
                     /note="matching EMBL:AA056416
                     RHdb:RH47592
                     dbSTS:STS40651
                     Identified using the e-PCR software (G. Schuler)"
     STS             177906. .178028
                     /note="matching EMBL:AA504155
                     RHdb:RH101916
                     dbSTS:STS69453
                     Identified using the e-PCR software (G. Schuler)"
     STS             177971. .178228
                     /note="matching EMBL:M63167
                     RHdb:RH69160
                     dbSTS:STS49050
                     Identified using the e-PCR software (G. Schuler)"
BASE COUNT    46690 a  60047 c  58929 g  41708 t
ORIGIN

Alignment Scores:
Pred. No.:              2.37e+03        Length:         207374
Score:                  92.00           Matches:        27
Percent Similarity:     50.00%          Conservative:   8
Best Local Similarity:  38.57%          Mismatches:     29
Query Match:            20.81%          Indels:         6

DB:                     9               Gaps:           3

US-09-724-000A-5 (1-81) x CNS07EFU (1-207374)

Qy       3 LeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePheSerThr 22
           |||||||||||||      ||||||   :::                        :::|||||||
Db 130256 CTGTTGGTTTTGGTTCCTTTGCTGGCTGTAGCCCACTAAACTGACTTCTCATGGAGCACA 130197

Qy      23 GluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHis 42
           |||    |||   :::|||                  |||:::                |||||
Db 130196 GAGGAAAAAGAAAAACCA------------GGCTCCAGGAAGGGCAGCTCCTCCTGTCAC 130149

Qy      43 ArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCysLys--- 61
           ||||||||||||||||   |||          :::|||       :::       |||:::
Db 130148 AGGGTCCCCAGTCCCACCTCTCCAGGCAGGAGGGGCGATGGTCTCACTTGCTGCAGATGA 130089

Qy      62 ProCysLysLeuGluProGluProArgLeu 71
           ||||||:::       |||   |||   |||
Db 130088 CCTTGTAGG---TGACCCCACCCAATACTC 130062


Search completed: May 17, 2003, 11:55:39
Job time : 2310 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:27:44 ; Search time 39 Seconds
(without alignments)
276.751 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLLVLSSLLCILLLCFSIF.....PCKLEPEPRLWVVPGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 229602

Minimum DB seq length: 23
Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	57	70.4	57	23	AAE16482	Human Secreted epi
2	30	37.0	30	23	AAE16485	Human Secreted epi
3	28	34.6	28	23	AAE16486	Human Secreted epi
4	23	28.4	23	23	AAE16484	Human Secreted epi
5	9	11.1	54	23	AAE16480	Mouse Secreted epi
6	9	11.1	77	23	AAE16483	Rat Secreted epith
7	9	11.1	78	23	AAE16479	Mouse Secreted epi
8	7	8.6	72	20	AAY19500	Amino acid sequenc
9	7	8.6	78	22	AAU45851	Propionibacterium
10	6	7.4	25	21	AAB39415	Human secreted pro
11	6	7.4	28	22	AAU05966	Cone snail O-supe
12	6	7.4	31	20	AAW81050	Signal peptide-cyt
13	6	7.4	32	22	AAM92181	Human digestive sy
14	6	7.4	39	4	AAP30235	Sequence of cortic
15	6	7.4	47	23	ABP41217	Human ovarian anti
16	6	7.4	48	21	AAB33769	Human secreted pro
17	6	7.4	52	21	AAG56482	Arabidopsis thalia
18	6	7.4	52	21	AAG58744	Arabidopsis thalia
19	6	7.4	52	22	AAU53929	Propionibacterium
20	6	7.4	52	22	AAU56623	Propionibacterium
21	6	7.4	53	23	ABP33245	Human ORF2218 prot
22	6	7.4	53	23	ABP34966	Human ORF3939 prot
23	6	7.4	54	20	AAY48477	Human breast tumou
24	6	7.4	54	22	AAM72480	Human bone marrow
25	6	7.4	54	22	AAM32728	Peptide #6765 enco
26	6	7.4	54	23	ABG42300	Human peptide enco
27	6	7.4	55	21	AAY91504	Human secreted pro
28	6	7.4	56	22	AAU55561	Propionibacterium
29	6	7.4	56	23	ABP07189	Human ORFX protein
30	6	7.4	58	22	AAM85937	Human immune/haema
31	6	7.4	59	23	ABP31870	Human ORF843 prote
32	6	7.4	65	22	AAU22012	Human cardiovascul
33	6	7.4	68	21	AAG09986	Arabidopsis thalia
34	6	7.4	69	22	ABB52742	Escherichia coli p
35	6	7.4	69	23	ABP09715	Human ORFX protein
36	6	7.4	70	21	AAB25064	Plant SDF encoded
37	6	7.4	72	22	ABB30289	Peptide #2940 enco
38	6	7.4	72	22	ABB35455	Peptide #2961 enco
39	6	7.4	72	22	ABB20895	Protein #2894 enco
40	6	7.4	72	22	AAM56278	Human brain expres
41	6	7.4	72	22	AAM68653	Human bone marrow
42	6	7.4	72	22	AAM16469	Peptide #2903 enco
43	6	7.4	72	22	AAM28962	Peptide #2999 enco
44	6	7.4	72	22	AAM04197	Peptide #2879 enco
45	6	7.4	72	23	ABG38233	Human peptide enco

ALIGNMENTS

RESULT 1
AAE16482
ID AAE16482 standard; peptide; 57 AA.
XX
AC AAE16482;
XX
DT 09-APR-2002 (first entry)
XX
DE Human Secreted epithelial colon stromal-1 (Secs-1) protein fragment.
XX
KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW human.
XX
OS Homo sapiens.
XX
PN WO200198497-A1.
XX
PD 27-DEC-2001.
XX
PF 28-NOV-2000; 2000WO-US32479.
XX
PR 21-JUN-2000; 2000US-0599087.
PR 28-NOV-2000; 2000US-0724000.
XX
PA (AMGE-) AMGEN INC.
XX
PI Polverino AJ, Luethy R;
XX

DR WPI; 2002-122281/16.
XX
PT Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT useful for diagnosing, treating and preventing hematopoietic disorder,
PT osteoporosis, Paget's disease, cancer, diabetes -
XX
PS Claim 14; Page 122; 134pp; English.
XX
CC The present invention relates to an isolated murine or human secreted
CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC therapy and cell therapy. Secs-1 is useful for identifying a compound
CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC preventing or ameliorating a disease condition such as haematopoietic
CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC is also useful for diagnosing a pathological condition which involves
CC determining the presence or amount of Secs-1 or polypeptide encoded by
CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC susceptibility to pathological condition based on the presence or amount
CC of expression of the polypeptide. The present sequence is human Secs-1
CC protein fragment.
XX
SQ Sequence 57 AA;

Query Match 70.4%; Score 57; DB 23; Length 57;
Best Local Similarity 100.0%; Pred. No. 6.9e-50;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPGALPQV 81
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 KRRPAKAWSGRRTRLCCHRVPSPNSTNLKGHHVRLCKPCKLEPEPRLWVVPGALPQV 57

RESULT 2
AAE16485
ID AAE16485 standard; peptide; 30 AA.
XX
AC AAE16485;
XX
DT 09-APR-2002 (first entry)
XX
DE Human Secreted epithelial colon stromal-1 (Secs-1) peptide #2.
XX
KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW human.
XX
OS Homo sapiens.
XX
PN WO200198497-A1.
XX
PD 27-DEC-2001.
XX
PF 28-NOV-2000; 2000WO-US32479.
XX
PR 21-JUN-2000; 2000US-0599087.
PR 28-NOV-2000; 2000US-0724000.
XX
PA (AMGE-) AMGEN INC.
XX
PI Polverino AJ, Luethy R;
XX
DR WPI; 2002-122281/16.
DR N-PSDB; AAD27026.
XX
PT Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT useful for diagnosing, treating and preventing hematopoietic disorder,
PT osteoporosis, Paget's disease, cancer, diabetes -
XX
PS Disclosure; Fig 4D; 134pp; English.
XX
CC The present invention relates to an isolated murine or human secreted
CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC therapy and cell therapy. Secs-1 is useful for identifying a compound
CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC preventing or ameliorating a disease condition such as haematopoietic
CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC is also useful for diagnosing a pathological condition which involves
CC determining the presence or amount of Secs-1 or polypeptide encoded by
CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC susceptibility to pathological condition based on the presence or amount
CC of expression of the polypeptide. The present sequence is human Secs-1
CC peptide encoded by second exon.
XX
SQ Sequence 30 AA;

Query Match 37.0%; Score 30; DB 23; Length 30;
Best Local Similarity 100.0%; Pred. No. 6.1e-23;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 GKRRPAKAWSGRRTRLCCHRVPSPNSTNLK 53
||||||||||||||||||||||||||||||
Db 1 GKRRPAKAWSGRRTRLCCHRVPSPNSTNLK 30

RESULT 3
AAE16486
ID AAE16486 standard; peptide; 28 AA.
XX
AC AAE16486;
XX
DT 09-APR-2002 (first entry)
XX
DE Human Secreted epithelial colon stromal-1 (Secs-1) peptide #3.
XX
KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW human.
XX
OS Homo sapiens.
XX
PN WO200198497-A1.
XX
PD 27-DEC-2001.
XX
PF 28-NOV-2000; 2000WO-US32479.
XX
PR 21-JUN-2000; 2000US-0599087.
PR 28-NOV-2000; 2000US-0724000.
XX
PA (AMGE-) AMGEN INC.
XX
PI Polverino AJ, Luethy R;
XX
DR WPI; 2002-122281/16.
DR N-PSDB; AAD27026.
XX
PT Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT useful for diagnosing, treating and preventing hematopoietic disorder,
PT osteoporosis, Paget's disease, cancer, diabetes -
XX
PS Disclosure; Fig 4F; 134pp; English.
XX

CC The present invention relates to an isolated murine or human secreted
CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC therapy and cell therapy. Secs-1 is useful for identifying a compound
CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC preventing or ameliorating a disease condition such as haematopoietic
CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC is also useful for diagnosing a pathological condition which involves
CC determining the presence or amount of Secs-1 or polypeptide encoded by
CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC susceptibility to pathological condition based on the presence or amount
CC of expression of the polypeptide. The present sequence is human Secs-1
CC peptide encoded by third exon.
XX
SQ Sequence 28 AA;

Query Match 34.6%; Score 28; DB 23; Length 28;
Best Local Similarity 100.0%; Pred. No. 6e-21;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 GHHVRLCKPCKLEPEPRLWVVPGALPQV 81
||||||||||||||||||||||||||||
Db 1 GHHVRLCKPCKLEPEPRLWVVPGALPQV 28

RESULT 4
AAE16484
ID AAE16484 standard; peptide; 23 AA.
XX
AC AAE16484;
XX
DT 09-APR-2002 (first entry)
XX
DE Human Secreted epithelial colon stromal-1 (Secs-1) peptide #1.
XX
KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW human.
XX
OS Homo sapiens.
XX
PN WO200198497-A1.
XX
PD 27-DEC-2001.
XX
PF 28-NOV-2000; 2000WO-US32479.
XX
PR 21-JUN-2000; 2000US-0599087.
PR 28-NOV-2000; 2000US-0724000.
XX
PA (AMGE-) AMGEN INC.
XX
PI Polverino AJ, Luethy R;
XX
DR WPI; 2002-122281/16.
DR N-PSDB; AAD27026.
XX
PT Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT useful for diagnosing, treating and preventing hematopoietic disorder,
PT osteoporosis, Paget's disease, cancer, diabetes -
XX
PS Disclosure; Fig 4A; 134pp; English.
XX
CC The present invention relates to an isolated murine or human secreted
CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC therapy and cell therapy. Secs-1 is useful for identifying a compound
CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC preventing or ameliorating a disease condition such as haematopoietic
CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC is also useful for diagnosing a pathological condition which involves
CC determining the presence or amount of Secs-1 or polypeptide encoded by
CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC susceptibility to pathological condition based on the presence or amount
CC of expression of the polypeptide. The present sequence is human Secs-1
CC peptide encoded by first exon.
XX
SQ Sequence 23 AA;

Query Match 28.4%; Score 23; DB 23; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.4e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRLLVLSSLLCILLLCFSIFSTE 23
|||||||||||||||||||||||
Db 1 MRLLVLSSLLCILLLCFSIFSTE 23

RESULT 5
AAE16480
ID AAE16480 standard; peptide; 54 AA.
XX
AC AAE16480;
XX
DT 09-APR-2002 (first entry)
XX
DE Mouse Secreted epithelial colon stromal-1 (Secs-1) protein fragment.
XX
KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW mouse.
XX
OS Mus musculus.
XX
PN WO200198497-A1.
XX
PD 27-DEC-2001.
XX
PF 28-NOV-2000; 2000WO-US32479.
XX
PR 21-JUN-2000; 2000US-0599087.
PR 28-NOV-2000; 2000US-0724000.
XX
PA (AMGE-) AMGEN INC.
XX
PI Polverino AJ, Luethy R;
XX
DR WPI; 2002-122281/16.
XX
PT Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT useful for diagnosing, treating and preventing hematopoietic disorder,
PT osteoporosis, Paget's disease, cancer, diabetes -
XX
PS Claim 14; Page 120; 134pp; English.
XX
CC The present invention relates to an isolated murine or human secreted
CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC therapy and cell therapy. Secs-1 is useful for identifying a compound
CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC preventing or ameliorating a disease condition such as haematopoietic
CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1

```
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is mouse Secs-1
CC   protein fragment.
XX
SQ   Sequence   54 AA;

  Query Match             11.1%;  Score 9;  DB 23;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      72 WVVPGALPQ 80
           |||||||||
Db      45 WVVPGALPQ 53


RESULT 6
AAE16483
ID   AAE16483 standard; Protein; 77 AA.
XX
AC   AAE16483;
XX
DT   09-APR-2002  (first entry)
XX
DE   Rat Secreted epithelial colon stromal-1 (Secs-1) protein.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   rat.
XX
OS   Rattus norvegicus.
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Disclosure; Fig 3; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is rat Secs-1
CC   protein.
XX
SQ   Sequence   77 AA;

  Query Match             11.1%;  Score 9;  DB 23;  Length 77;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      72 WVVPGALPQ 80
           |||||||||
Db      68 WVVPGALPQ 76


RESULT 7
AAE16479
ID   AAE16479 standard; Protein; 78 AA.
XX
AC   AAE16479;
XX
DT   09-APR-2002  (first entry)
XX
DE   Mouse Secreted epithelial colon stromal-1 (Secs-1) protein.
XX
KW   Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
KW   haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
KW   Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
KW   diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
KW   anorectic; immunomodulator; antipsoriatic; vulnerary; antiinfertility;
KW   gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
KW   mouse.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..24
FT                   /label= Signal_peptide
FT   Protein         25..78
FT                   /label= Mature_mouse_Secs-1_protein
XX
PN   WO200198497-A1.
XX
PD   27-DEC-2001.
XX
PF   28-NOV-2000; 2000WO-US32479.
XX
PR   21-JUN-2000; 2000US-0599087.
PR   28-NOV-2000; 2000US-0724000.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Polverino AJ,  Luethy R;
XX
DR   WPI; 2002-122281/16.
DR   N-PSDB; AAD27024.
XX
PT   Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT   useful for diagnosing, treating and preventing hematopoietic disorder,
PT   osteoporosis, Paget's disease, cancer, diabetes -
XX
PS   Claim 13; Fig 3; 134pp; English.
XX
CC   The present invention relates to an isolated murine or human secreted
CC   epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC   variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC   therapy and cell therapy. Secs-1 is useful for identifying a compound
CC   which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC   preventing or ameliorating a disease condition such as haematopoietic
CC   disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC   disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC   chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC   is also useful for diagnosing a pathological condition which involves
CC   determining the presence or amount of Secs-1 or polypeptide encoded by
CC   Secs-1 DNA in a sample; and diagnosing a pathological condition, or
```

```
CC   susceptibility to pathological condition based on the presence or amount
CC   of expression of the polypeptide. The present sequence is mouse Secs-1
CC   protein.
XX
SQ   Sequence   78 AA;

  Query Match             11.1%;  Score 9;  DB 23;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      72 WVVPGALPQ 80
           |||||||||
Db      69 WVVPGALPQ 77


RESULT 8
AAY19500
ID   AAY19500 standard; Protein; 72 AA.
XX
AC   AAY19500;
XX
DT   14-JUL-1999  (first entry)
XX
DE   Amino acid sequence of a human secreted protein.
XX
KW   Human secreted protein; cancer; tumour; neurodegenerative disorder;
KW   developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW   immune system disease; autoimmune disease; hepatic disease;
KW   renal disease; lymphoma; inflammation; allergy; ischemic shock;
KW   Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
KW   obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
KW   lung disease; thymus disease; digestive disorder; endocrine disorder;
KW   infection; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO9922243-A1.
XX
PD   06-MAY-1999.
XX
PF   23-OCT-1998;   98WO-US22376.
XX
PR   24-OCT-1997;   97US-0063387.
PR   24-OCT-1997;   97US-0062784.
PR   24-OCT-1997;   97US-0063088.
PR   24-OCT-1997;   97US-0063089.
PR   24-OCT-1997;   97US-0063090.
PR   24-OCT-1997;   97US-0063091.
PR   24-OCT-1997;   97US-0063092.
PR   24-OCT-1997;   97US-0063097.
PR   24-OCT-1997;   97US-0063098.
PR   24-OCT-1997;   97US-0063099.
PR   24-OCT-1997;   97US-0063100.
PR   24-OCT-1997;   97US-0063101.
PR   24-OCT-1997;   97US-0063109.
PR   24-OCT-1997;   97US-0063110.
PR   24-OCT-1997;   97US-0063111.
PR   24-OCT-1997;   97US-0063148.
PR   24-OCT-1997;   97US-0063386.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Brewer LA,  Carter KC,  Duan DR,  Ebner R,  Endress GA;
PI   Feng P,  Florence C, Florence KA,  Greene JM,  Janat F;
PI   Kayw H,  Lafleur DW,  Moore PA,  Ni J, Olsen HS,  Rosen CA;
PI   Ruben SM,  Shi Y,  Soppet DR,  Wei Y,  Young P;
XX
DR   WPI; 1999-303069/25.
DR   N-PSDB; AAX61380.
XX
PT   New isolated human genes and the secreted polypeptides they encode
XX
PS   Claim 11; Page 416; 546pp; English.
XX
CC   The specification describes cDNA sequences (AAX61322-X61470) encoding
CC   human secreted proteins (AAY19442-Y19590). The polynucleotides and their
CC   corresponding secreted polypeptides are useful for preventing, treating
CC   or ameliorating medical conditions, e.g. by protein or gene therapy.
CC   Pathological conditions can also be diagnosed by determining the amount
CC   of the polypeptides in a sample or by determining the presence of
CC   mutations in the polynucleotides. Specific uses are described for each
CC   of the polynucleotides, based on which tissues they are most highly
CC   expressed in, and include developing products for the diagnosis or
CC   treatment of cancer, tumours, neurodegenerative disorders, developmental
CC   abnormalities and fetal deficiencies, blood disorders, leukemias,
CC   diseases of the immune system, autoimmune diseases, hepatic and renal
CC   disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
CC   and cognitive disorders, schizophrenia, prostate diseases, obesity,
CC   disorders involving osteoclasts such as osteoporosis, arthritis or
CC   malignancies, diseases of testes, lung or thymus, digestive/endocrine
CC   disorders, infections and AIDS. The polypeptides are also useful for
CC   identifying their binding partners.
XX
SQ   Sequence   72 AA;

  Query Match             8.6%;  Score 7;  DB 20;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 LLLCFSI 19
           |||||||
Db      22 LLLCFSI 28


RESULT 9
AAU45851
ID   AAU45851 standard; Protein; 78 AA.
XX
AC   AAU45851;
XX
DT   27-FEB-2002  (first entry)
XX
DE   Propionibacterium acnes immunogenic protein #6747.
XX
KW   SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW   dermatological; osteopathic; neuroprotectant.
XX
OS   Propionibacterium acnes.
XX
PN   WO200181581-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US12865.
XX
PR   21-APR-2000; 2000US-199047P.
PR   02-JUN-2000; 2000US-208841P.
PR   07-JUL-2000; 2000US-216747P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Persing DH,  Mitcham JL,  Wang SS,  Bhatia A;
PI   L'maisonneuve J,  Zhang Y,  Jen S,  Carter D;
XX
DR   WPI; 2001-616774/71.
DR   N-PSDB; AAS59528.
XX
PT   Propionibacterium acnes polypeptides and nucleic acids useful for
PT   vaccinating against and diagnosing infections, especially useful for
PT   treating acne vulgaris -
XX
PS   Example 1; SEQ ID No 7046; 1069pp; English.
XX
CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
```

```
CC   polypeptides. The proteins and their associated DNA sequences are used in
CC   the treatment, prevention and diagnosis of medical conditions caused by
CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC   pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC   P. acnes is also involved in infections of bone, joints and the central
CC   nervous system, however it is particularly involved in the inflammatory
CC   lesions associated with acne vulgaris. A method for detecting the
CC   presence or absence of P. acnes in a patient comprises contacting a
CC   sample with a binding agent that binds to the proteins of the invention
CC   and determining the amount of bound protein in the sample. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   specific for P. acnes proteins. These antibodies can be used to
CC   downregulate expression and activity of P. acnes polypeptides and
CC   therefore treat P. acnes infections. The antibodies may also be used as
CC   diagnostic agents for determining P. acnes presence, for example, by
CC   enzyme linked immunosorbent assay (ELISA).
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   78 AA;

  Query Match             8.6%;  Score 7;  DB 22;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      30 KAWSGRR 36
           |||||||
Db      55 KAWSGRR 61


RESULT 10
AAB39415
ID   AAB39415 standard; Protein; 25 AA.
XX
AC   AAB39415;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Human secreted protein sequence encoded by gene 14 SEQ ID NO:74.
XX
KW   Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW   nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW   antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
KW   food additive; preservative; chromosome identification; cancer;
KW   female reproductive system disorder; immune disorder; wound healing;
KW   cardiovascular disorder; neurological disease; infectious disease;
KW   infection.
XX
OS   Homo sapiens.
XX
PN   WO200058340-A2.
XX
PD   05-OCT-2000.
XX
PF   23-MAR-2000; 2000WO-US07724.
XX
PR   26-MAR-1999;   99US-0126510.
PR   07-JAN-2000; 2000US-0174850.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-594638/56.
DR   N-PSDB; AAC74350.
XX
PT   Fifty nucleic acid molecules encoding human secreted proteins, useful
PT   in the prevention, treatment and diagnosis of cancer, immune disorders,
PT   cardiovascular disorders and neurological diseases -
XX
PS   Claim 11; Page 362; 391pp; English.
XX
CC   The polynucleotide sequences given in AAC74337 to AAC74386 encode the
CC   human secreted proteins given in AAB39402 to AAB39451. AAB39452 to
CC   AAB39484 represent human secreted polypeptide sequences and proteins
CC   homologous to them, which are given in the exemplification of the present
CC   invention. Human secreted proteins have activities based on the tissues
CC   and cells the genes are expressed in. Example of activities include:
CC   cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
CC   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
CC   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC   cardiant. The polynucleotides and polypeptides are useful for preventing,
CC   treating or ameliorating a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
CC   can also be used as a food additive or preservative to increase or
CC   decrease storage capabilities. The polynucleotide are useful for
CC   chromosome identification. They are also useful as probes for diagnosing
CC   a disorder related to the female reproductive system, particularly breast
CC   and/or ovary cancer. They are also useful in the gene therapy of breast
CC   and ovarian cancer. Secreted protein nucleic acids, proteins,
CC   antibodies, agonists and antagonists are useful in the diagnosis,
CC   treatment and prevention of: (a) cancer; (b) immune disorders; (c)
CC   cardiovascular disorders; (d) wound healing; (e) neurological diseases;
CC   and (f) infectious diseases such as viral, bacterial, fungal and
CC   parasitic infections. AAC74328 to AAC74336 and AAB39401 represent
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   25 AA;

  Query Match             7.4%;  Score 6;  DB 21;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 79;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 VLSSLL 10
            ||||||
Db      11 VLSSLL 16


RESULT 11
AAU05966
ID   AAU05966 standard; Peptide; 28 AA.
XX
AC   AAU05966;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Cone snail  O-superfamily conotoxin, De6.2.
XX
KW   Cone snail; O-superfamily conotoxin; sodium channel;
KW   demyelinating disease; multiple sclerosis; Huntingdon's disease;
KW   neuropathy; carpal tunnel syndrome; cardiovascular disorder;
KW   congestive heart failure; cancer; immunosuppression; epilepsy;
KW   asthma; ischaemia; stroke; pain.
XX
OS   Conus delessertii.
XX
PN   WO200149312-A2.
XX
PD   12-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35431.
XX
PR   30-DEC-1999;   99US-0173754.
PR   26-JUN-2000; 2000US-0214263.
PR   20-JUL-2000; 2000US-0219440.
PR   27-OCT-2000; 2000US-0243412.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
PA   (COGN-) COGNETIX INC.
XX
PI   Olivera BM,  Cartier GE,  Watkins M,  Hillyard DR,  McIntosh JM;
PI   Layer RT,  Jones RM;
XX
DR   WPI; 2001-418352/44.
```

```
XX
PT   New O-superfamily polypeptides useful for treating voltage gated ion
PT   channel disorders, including demyelinating diseases i.e. multiple
PT   sclerosis -
XX
PS   Claim 2; Page 69; 277pp; English.
XX
CC   The sequence is a cone snail O-superfamily conotoxin peptide.
CC   The peptides are useful for regulating the flow of sodium through
CC   sodium channels in an individual and the treatment or prevention of
CC   disorders associated with voltage gated ion channel disorders,
CC   including demyelinating diseases i.e. multiple sclerosis, optic
CC   neuromyelitis, disseminated encephalomyelitis, adrenoleukodystrophy,
CC   acute transverse myelitis, progressive multifocal leukoencephalopathy,
CC   sub acute sclerosing panecephalomyelitis (SSPE), metachromatic
CC   leukodystrophy, Pelizaeus-Merzbacher disease, spinal cord injury,
CC   botulinum toxin poisoning, Huntington's, compression, entrapment
CC   neuropathies i.e. ulnar nerve palsy, and carpal tunnel syndrome,
CC   cardiovascular disorders, i.e. cardiac arrhythmias and congestive heart
CC   failure, reactive gliosis, hyperglycaemia, immunosuppression, cocaine
CC   addiction, cancer, cognitive dysfunction, neurotransmitter disorders
CC   (i.e. Eaton-Lambert syndrome) and reversal of curare and other
CC   neuromuscular blocking drugs. The neurological disorder is a seizure,
CC   preferably one associated with epilepsy. The neurological disorder is a
CC   neurotoxic injury associated with hypoxia, anoxia or ischaemia. The
CC   neurotoxic injury is associated with stroke, cerebrovascular accident,
CC   brain or spinal cord trauma, myocardial infarct, physical trauma,
CC   drownings, suffocation, perinatal asphyxia or hypoglycaemic events. The
CC   disorder is pain i.e. migraine, acute pain, persistent pain,
CC   neuropathic pain, nociceptive pain. The disorder is inflammation or a
CC   cardiovascular disorder. A conotoxin peptide of is useful to
CC   alleviate pain in a mammal in pain or about to be subjected to a pain
CC   causing event, and to treat disorders associated with radical
CC   depolarisation of excitable membranes by activating a KATP channel, the
CC   disorders include cardiac, ocular and cerebral ischaemia and asthma.
XX
SQ   Sequence   28 AA;

  Query Match             7.4%;  Score 6;  DB 22;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 CILLLC 16
           ||||||
Db      22 CILLLC 27


RESULT 12
AAW81050
ID   AAW81050 standard; Peptide; 31 AA.
XX
AC   AAW81050;
XX
DT   10-MAY-1999  (first entry)
XX
DE   Signal peptide-cytotoxic T cell epitope fusion.
XX
KW   MG50; melanoma gene-50; melanoma associated antigen; human;
KW   cancer; lung cancer; rhabdomyosarcoma; diagnosis; therapy; vaccine;
KW   cytotoxic T cell epitope; signal peptide.
XX
OS   Chimeric - Synthetic.
OS   Chimeric - Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..21
FT                   /note= "signal peptide"
FT   Peptide         22..31
FT                   /note= "cytotoxic T cell epitope"
XX
PN   WO9855133-A1.
XX
PD   10-DEC-1998.
XX
PF   04-JUN-1998;   98WO-US11533.
XX
PR   06-JUN-1997;   97US-0870941.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYSC-) UNIV SOUTHERN CALIFORNIA.
XX
PI   Deans RJ,  Kan-Mitchell J,  Minev BR,  Mitchell MS;
XX
DR   WPI; 1999-080820/07.
XX
PT   New MG50 melanoma associated antigen fragments - used to develop
PT   products for the detection, treatment and prevention of
PT   MG50-expressing cancers, e.g. melanoma, lung cancer or
PT   rhabdomyosarcoma
XX
PS   Claim 12; Page 58; 79pp; English.
XX
CC   This peptide comprises a claimed cytotoxic T cell epitope composed
CC   of a signal peptide (see AAW81047) and a T cell epitope (see AAW81032)
CC   derived from the human MG50 melanoma associated antigen (see
CC   AAW81030). The signal peptide can facilitate presentation of the
CC   epitope as a complex with an MHC molecule at the surface of an
CC   antigen presenting cell (APC). The invention provides additional
CC   T cell epitopes (see AAW81031-45) from MG50, and APCs that can
CC   express an MG50 T cell epitope complexed with an MHC molecule on
CC   their surfaces, and which can be administered to a patient having
CC   a cancer that contains MG50-expressing cells in order to stimulate
CC   an active immune response. MG50 vaccines comprising an MG50
CC   polypeptide, a T cell epitope optionally attached to a signal
CC   peptide, an anti-idiotypic antibody or a nucleic acid molecule
CC   encoding an MG50 polypeptide or T cell epitope can be administered
CC   for preventative or therapeutic purposes. The products and methods
CC   are used for the detection, treatment and prevention of e.g.
CC   melanoma, lung cancer and rhabdomyosarcoma.
XX
SQ   Sequence   31 AA;

  Query Match             7.4%;  Score 6;  DB 20;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 LLLCFS 18
           ||||||
Db      11 LLLCFS 16


RESULT 13
AAM92181
ID   AAM92181 standard; Protein; 32 AA.
XX
AC   AAM92181;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human digestive system antigen SEQ ID NO: 1530.
XX
KW   Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW   ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW   digestive system disorder; Meckel's diverticulum.
XX
OS   Homo sapiens.
XX
PN   WO200155314-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01324.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
```

PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-502630/55.
DR N-PSDB; AAK87954.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognozing disorders of the
PT digestive system, particularly cancer and cancer metastases -
XX
PS Claim 11; SEQ ID NO 1530; 986pp; English.

```
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human digestive system antigens. These can be used in the
CC   diagnosis, treatment and prevention of digestive system disorders,
CC   including cancer, Meckel's diverticulum, bacterial or parasitic
CC   infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC   ulcerative colitis. The present sequence is a digestive system antigen of
CC   the invention.
XX
SQ   Sequence   32 AA;

  Query Match             7.4%;  Score 6;  DB 22;  Length 32;
  Best Local Similarity   100.0%;  Pred. No. 99;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 LLLCFS 18
           ||||||
Db      13 LLLCFS 18


RESULT 14
AAP30235
ID   AAP30235 standard; peptide; 39 AA.
XX
AC   AAP30235;
XX
DT   25-APR-1992  (first entry)
XX
DE   Sequence of corticotropin, adrenocorticotripic hormone (ACTH)
DE   analogue Tyr23-Phe2-Nle4-ACTH.
XX
KW   Hormone; pituitary; growth; adrenal gland.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= H-Ser
FT   Misc-difference 3
FT                   /label= Nle
FT   Misc-difference 19
FT                   /label= Nal
FT   Modified-site   39
FT                   /label= Phe-OH
XX
PN   US4415546-A.
XX
PD   15-NOV-1983.
XX
PF   16-MAY-1983;   83US-0302746.
XX
PR   12-MAY-1981;   81US-0262974.
XX
PA   (RAMA/) RAMACHANDRAN J.
XX
PI   Ramachandran J,  Buckley DL,  Yamashiro DH,  Hagman JR;
XX
DR   WPI; 1983-832889/48.
XX
PT   ACTH analogues and their radio:iodinated derivs. - useful as
PT   corticosterone stimulators and in radioimmunoassays
XX
PS   Claim 1; Col 10; 6pp; English.
XX
CC   The ACTH analogues have biological activity fully comparable to that
CC   of native ACTH and so are useful for stimulation of corticosterones.
CC   The radioiodinated derivs. are useful in radioimmunoassays, and in
CC   investigations of body processes, and unlike radioiodinated ACTH,
CC   they do not show activity loss.
XX
SQ   Sequence   39 AA;

  Query Match             7.4%;  Score 6;  DB 4;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      25 KRRPAK 30
           ||||||
Db      16 KRRPAK 21


RESULT 15
ABP41217
ID   ABP41217 standard; Protein; 47 AA.
XX
AC   ABP41217;
XX
DT   23-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HOCPO31, SEQ ID NO:2349.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US18569.
XX
PR   07-JUN-2000; 2000US-209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   N-PSDB; ABQ54294.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g.
PT   ovarian cancer), immune disorders, cardiovascular disorders and
PT   neurological diseases -
XX
PS   Claim 11; SEQ ID No 2349; 2922pp; English.
XX
CC   The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC   encompasses polypeptides 90% identical and polynucleotides 95% identical
CC   to the sequences of the invention. The invention additionally relates to
CC   recombinant vectors and host cells comprising human ovarian antigen
CC   polynucleotides, antibodies against human ovarian antigens, and the use
CC   of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC   treating, prognosing or preventing various ovary and/or breast-related
CC   disorders. Such conditions include ovarian cancer and breast cancer, and
CC   metastatic tumours of ovarian or breast origin, reproductive system
CC   disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC   polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC   disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC   shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC   vaginitis), immune disorders (e.g., congenital and acquired
CC   immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC   blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC   respiratory disorders, neurological disorders, gastrointestinal disorders
CC   and urinary system disorders. Ovarian antigen polypeptides and
CC   polynucleotides may also be used in screening for compounds which
```

CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 47 AA;

Query Match 7.4%; Score 6; DB 23; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLLC 11
||||||
Db 6 LSSLLC 11

Search completed: May 16, 2003, 15:34:06
Job time : 41 secs

73 x 2